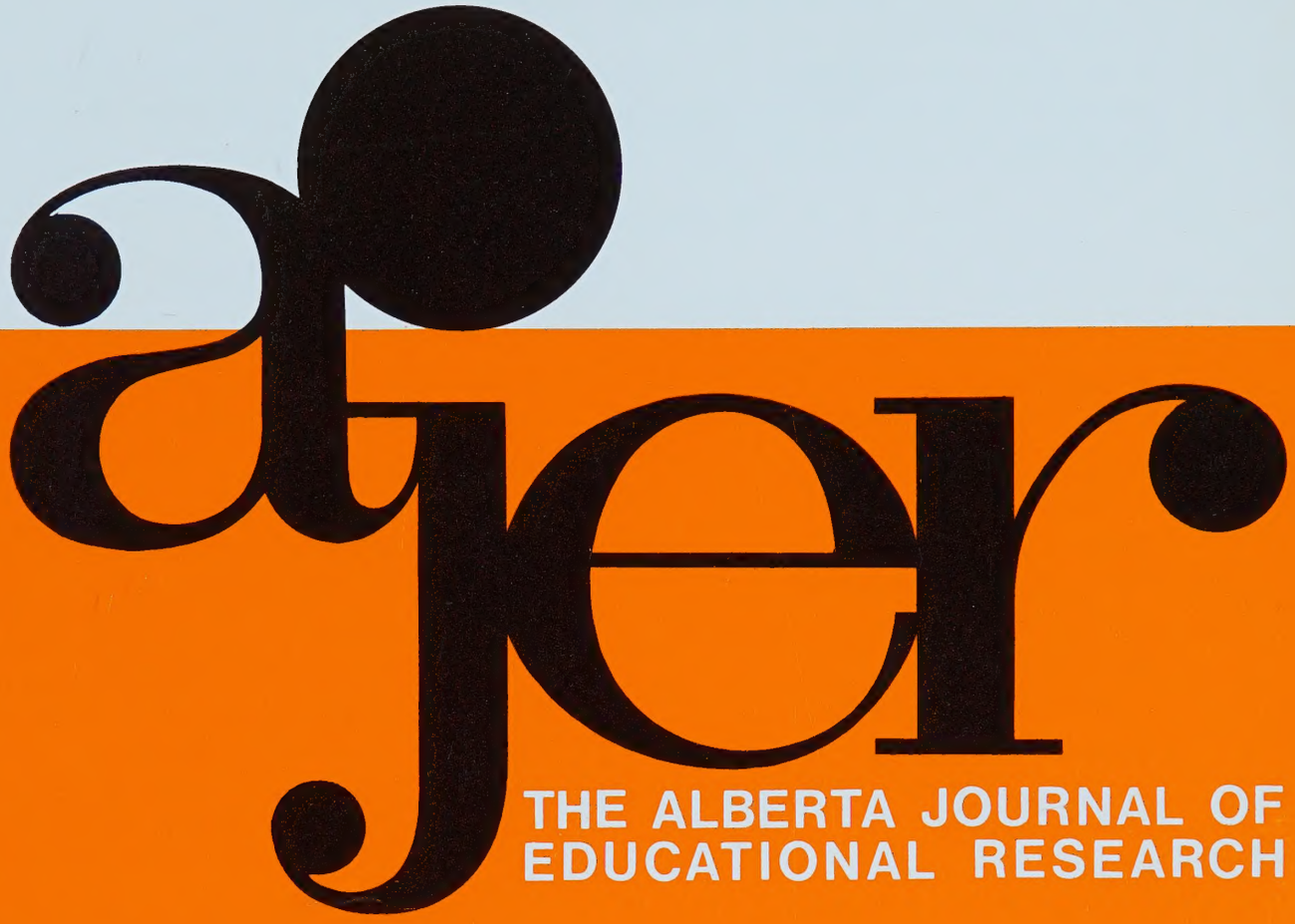

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

VOLUME XXXII NUMBER 1 MARCH 1986

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONTON

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, and December by the
Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: Walter H. Worth

CONSULTING EDITORS

Reno A. Bosetti,
Alberta Department of Education

Robert C. Bryce,
University of Regina

Robert K. Crocker,
Memorial University of Newfoundland

Kathy G. Dueck,
University of Calgary

Michael G. Fullan,
Ontario Institute for Studies in Education

Myrna L. Greene,
University of Lethbridge

Geoffrey B. Isherwood,
McGill University

Joanne C. Keselman,
University of Manitoba

Terry R. Morrison,
University of Athabasca

Alan R. Pence,
University of Victoria

Bikkar S. Randhawa,
University of Saskatchewan

Jim Richardson,
University of New Brunswick

W. Todd Rogers,
University of British Columbia

Jaap Tuinman,
Simon Fraser University

Donald J. Willower,
Pennsylvania State University

John D. Wilson,
Moray House College of Education

FACULTY ADVISORY COMMITTEE

Thomas O. Maguire

Erwin Miklos (Chair)

Alton T. Olson

Daiyo Sawada

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

The subscription rate is $24.00 per year for individuals, $32.00 per year for institutions. Add $7.00 for delivery outside of Canada. Single copies are $7.00 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXXII, Number 1 *March, 1986*

CONTENTS

THE EDITOR'S PAGE: *Integrating SSHRC's New Programs* 1

Predictive Validity of the Canadian Readiness Test and Selected Measures of Cognitive Development for Grade One Reading and Mathematics Achievement 2
B.Collis, L.O. Ollila, and L.D. Yore

An Analysis of WISC-R Item ResponsesWith Canadian Inuit Children 12
H.H. Mueller, R.F. Mulcahy, L. Wilgosh, B. Watters, and G.J. Mancini

Using Secondary Analysis of National Assessment Data to Identify Predictors of Junior High School Students' Outcomes 37
B.J. Fraser, W.W. Welch, and H.J. Walberg

Student Orientations to School Rules 51
A.J. Macdonell and W.B.W. Martin

Benefits and Disappointments from Participation in the Alberta High School Equivalency Program 66
B.C. Carbol and T.O. Maguire

BOOK REVIEWS

Curriculum Canada V., D.A. Roberts and J.O. Fritz, Eds. 77
Reviewed by J.H. Young

FACULTY OF EDUCATION
The University of Alberta

ACKNOWLEDGMENT

The review process is central to the quality of this journal. We thank the following people who reviewed manuscripts in 1985. Their thoughtful critiques contribute to the advancement of theory, research, and practice in education.

Armstrong, R.D.
Asp, E.
Assheton-Smith, M.I.
Atherton, P.J.
Baine, D.A.
Bjorkquist, D.C.
Blakey, J.M.
Blase, J.J.
Blinkhorn, S.
Boorman, J.
Bryant, D.E.
Bulcock, J.
Bumbarger, C.S.
Campbell, C.E.
Campbell, D.S.
Carey, S.T.
Carson, T.R.
Cathcart, W.G.
Clandinan, D.J.
Coleman, P.
Collis, B.
Connelly, M.
Cordis, L.L.
Cote, A.
Crocker, R.K.
Das, J.P.
De Faveri, I.
Downey, L.W.
Dueck, K.G.
Edwards, J.P.
Evans, P.O.
Fitzgerald, D.
Fitzsimmons, G.W.
Freebody, P.
Freeman, P.
Frender, R.H.
Friesen, D.
Grant, J.
Grimmett, P.P.
Gue, L.R.
Heffernan, P.J.
Henchey, N.
Herson, N.L.
Hill, R.J.
Holdaway, E.A.
Housego, B.E.J.
Housego, I.E.
Ingram, E.J.
Isherwood, G.B.
Jackson, R.K.
Johnson, F.T.
Kach, N.
Kass, H.J.
Keselman, J.C.
Kieren, T.E.
Knox, G.
Lapkin, S.J.
Lee, H.D.
Liedtke, W.W.
Lucas, B.
MacConnell, F.M.
Maguire, T.O.
Malicky, G.V.
Marino, J.G.
Martin, M.
Martin, W.B.W.
McLean, G.
McLean, L.D.
Mickelson, N.
Mills, J.
Morrison, T.R.
Muir, N.
Murtagh, A.
Nyberg, V.R.
Ott, W.
Parsons, J.B.
Pawlovich, W.E.
Peters, H.O.
Pollard, M.J.
Pomfret, A.
Proudfoot, A.
Puffer, K.
Randhawa, B.S.
Ratsoy, E.W.
Richards, D.M.
Richardson, C.J.
Robertaille, D.F.
Robertson, J.E.
Rogers, W.T.
Romaniuk, E.W.
Russell, J.
Sackney, L.E.
Saklofske, D.H.
Sande,D.J.
Schroeder, T.
Seger, J.E.
Seifert, K.L.
Sitko, M.C.
Toews, W.
Ubelacker, S.D.
Unruh, W.W.
Violato, C.
Walker, L.
Wangler, D.G.
Ward, K.L.
Watson, R.J.
Willower, D.J.
Wilson, R.J.
Woodley, K.F.
Young, J.H.

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 1-1*

THE EDITOR'S PAGE

Integrating SSHRC's New Programs

As this is being written, members of parliament and cabinet ministers are being subjected to intensive lobbying in support of the Social Sciences and Humanities Research Council's *Five-Year Plan for Financing Research in the Social Sciences and Humanities, 1985-1990.* Three features of the plan could have a substantial impact upon the Canadian educational research community.

One is the inclusion of "education for the post-industrial society" among three new themes for targeted research deemed to be in the national interest. Research related to this theme is intended to help individual Canadians, and their educational system, better meet the challenges of rapidly changing technology and market patterns.

A second feature is the introduction of five-year "Canada Research Fellowships" to stimulate research activity immediately, and to provide a partial solution to the looming problem of staffing advanced education and research in the 1990s.

The third is the creation of "research centres" to focus research on persisting national problems, promote interdisciplinary and multidisciplinary research, provide research training, offer facilities for visiting researchers and the proposed Canada research fellows, act as nodes in communication networks, and link researchers and research users in government and the private sector.

As this is being read, some of the decisions about the funding of the plan will have already been made by the federal government and the Council. The case in support of the three new features just enumerated is so compelling that it would be as surprising as it would be disastrous for them not to be pursued in some fashion.

Equally compelling is the need to integrate research on education for the post-industrial society with the Canada research fellowships in a number of research centres. That integration maximizes impact has been demonstrated in connection with the study of Canada's aging population during SSHRC's first five-year plan. But integration does not just happen. It must be deliberately nurtured.

Initially at least, nurturing the integration of SSHRC's new programs in education is a leadership opportunity and obligation of deans of faculties of education. Theirs is a key role. They head the organization with the largest pool of relevant research talent. They are influential in shaping expectations for the use of that talent. They are in a position to communicate effectively with the broader university community, provincial governments, teachers' and trustees' organizations, and local school districts. Thus, each of them is in a position to be the "powerful sponsor" and "continuing advocate" that has characterized every successful planning endeavor of the past in Canadian education.

If we had a "nurturer of the year" award in educational research in Canada for 1986, I'd bet the recipient would be a dean of education.

WHW

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 1-1*

BETTY COLLIS

LLOYD O. OLLILA

and

LARRY D. YORE
University of Victoria

Predictive Validity of the Canadian Readiness Test and Selected Measures of Cognitive Development for Grade One Reading and Mathematics Achievement

Correlational associations of the Canadian Readiness Test (CRT) and selected measures of cognitive development with subsequent achievement in reading and mathematics were examined. Four subscales of the CRT were identified as being highly involved as predictors. Conservation of number demonstrated some potential as a predictor of future achievement, but overall the addition of scores from Piagetian-type tasks to CRT data contributed little to the prediction of either reading or mathematics achievement. The relationship between the constructs underlying the CRT and the cognitive development tasks was also examined.

Prediction prior to Grade One of children's readiness for reading and mathematics instruction and of their subsequent achievement in these areas is important for both practical and theoretical reasons. The practical concerns focus on the efficient and effective prediction of subsequent achievement using a concise, valid and reliable battery of tests. One of the purposes of this study is to further validate an existing readiness test in order to strengthen support for its use as an achievement predictor. Readiness testing is of interest from a theoretical perspective for at least two reasons: to aid in the identification of underlying skills and processes which are fundamental to learning how to read and how to compute mathematically, and also for the investigation of the relationship between cognitive development defined

---

Dr. Collis is an Assistant Professor in the Department of Psychological Foundations, University of Victoria, with research interests in measurement and evaluation of primary students, and computer applications in education.

Dr. Ollila is a Professor in the Department of Communication and Social Foundations, University of Victoria, with research interests in primary language arts.

Dr. Yore is an Associate Professor in the Department of Social and Natural Science, University of Victoria, with research interests in cognitive development and science education.

within a Piagetian framework and early school achievement. A second purpose of this study involves both these perspectives and focuses on the gain in theoretical knowledge and empirical prediction which may occur when readiness test results are supplemented with information derived from the administration of selected Piagetian tasks.

The Canadian Readiness Test as a Predictor of Achievement.

Studies by Gates, Bond and Russell (1939); Smith (1957); Barrett (1965); Bond and Dykstra (1967); and Johnson (1969) indicated predictive correlations between comprehension of oral language, usage of oral language, visual discrimination, auditory discrimination, knowledge of quantitative relationships, shape completion, sensorimotor skills related to handwriting, and following directions with later reading achievement. The Metropolitan Readiness Test (Hildreth, et al., 1965) and the Clymer-Barrett Prereading Battery (Clymer & Barrett, 1967) were designed around these results. However, Evanechko, Ollila, Downing, and Braun (1973) suggested a broader base was needed for effective readiness testing. They noted that:

> Less attention has been given to the non-perceptual components which constitute part of the total complex of reading readiness skills. Technical linguistic terms such as "word," "letter," "number," "sound," etc., are often used by teachers when talking to their first grade students. Similarly, teachers assume that children possess the ability to conceptualize and label real and vicarious experiences, that is, that they possess the ability to associate meaning with verbal symbols. They also assume the child's syntactical and morphological competence sufficient to enable him to form and comprehend grammatical structure in language. Further, the child's cognitive functioning is infrequently measured as an indicator of the ability to cope with the memory requirements of beginning reading. Similarly the rate at which the child can assimilate the visual linguistic stimuli is seldom considered by those predicting reading success. (p. 62)

Based on this analysis, the Canadian Readiness Test (Evanechko, et. al., 1973), a battery of seven subscales, was developed to measure four general areas of prereading behavior—the child's concept of the reading task, perceptual ability, linguistic competence, and level of cognitive functioning. In the original study, perceptual ability appeared to be the best predictor of reading achievement, followed closely by cognitive functioning ability, linguistic ability, and concept of the reading task. It was argued that subscales relating to each of these areas were important in strengthening both the diagnostic and predictive validity of reading readiness tests. In particular it was noted that word recognition behaviors were dependent upon the child's readiness in letter recognition, learning rate, listening and semantics; comprehension of ideas depended upon readiness in letter recognition, semantics, and morphology; comprehension of instructions depended upon readiness in letter recognition, word matching, syntax, orientation of literacy, and ending sounds. Total achievement in reading was best predicted by the child's readiness as measured by the subscales Letter Recognition, Semantics, and Beginning Sounds. The authors concluded that "to serve a diagnostic function, a readiness test should have a range of subtests including items which ensure the child's concept of the reading task, his perceptual ability, linguistic competence, and cognitive functioning" (p. 78).

Cognitive Development as a Predictor of Achievement

Various researchers have explored cognitive development and related Piagetian tasks as predictors of reading and mathematics achievement (Waller, 1977).

Kingma and Koops (1983) found Piagetian tasks that measured seriation, conservation and multiplicative classification to be equal or superior to traditional intelligence tests in predicting mathematics achievement. Taloumis (1979) found scores on conservation and measurement tasks to be highly predictive of mathematics achievement, and Dimitrovsky and Almy (1975) indicated early achievement of conservation ability to be predictive of later reading performance. Jordan and Brownlee (1981) reported results of a meta-analysis examining the relationship between cognitive development and school achievement tests and indicated that Piagetian tasks account for approximately 29% of the variance in mathematics achievement and 16% of the variance in reading achievement within the studies involved in the meta-analysis.

However, other researchers have not found Piagetian tasks to be effective readiness predictors. Brekke and Williams (1975) for example, found conservation not to be a significant predictor of reading achievement. Nelson (1980) presented a similar conclusion about the predictive nature of cognitive functioning as measured by Piagetian tasks with respect to mathematics achievement. He summarized research in this area by noting:

> Ten years of research have failed to establish any practical connection between conservation and mathematics learning. What is an essential and central concept in Piaget's developmental psychology seems to have little bearing on how children learn mathematics. (1980, p. xi)

Arlin (1981) suggested that the lack of consistent evidence relating cognitive development tasks and achievement may be a result of the use of an insufficient or inappropriate battery of Piagetian tasks "with little rationale provided for the selection of one Piagetian task in preference to another" (p. 712). She found the three subsystems of Piagetian tasks—classification, seriation, and conservation—"together appear to make significant contributions to a child's achievement in reading and mathematics . . [and] that knowledge of a child's general concrete operativity may be an important component of our understanding of the child's general readiness for mathematics and reading instruction." (p. 270)

Against this background, the present study examines the following questions:

1. Is the CRT an effective predictor of achievement at the end of Grade One?
2. To what extent are the subscales of the CRT (and consequently its underlying theoretical model) involved in subsequent school achievement?
3. Does the addition of selected Piagetian tasks to the CRT battery improve prediction of Grade One achievement?
4. What is the relationship between the constructs underlying the CRT and the cognitive development constructs represented by the Piagetian task battery? Is this relationship pertinent to advancing what we know about how children learn?

Method

Sample

A sample of 99 middle-class, mixed racial, English-speaking children enrolled in four Grade One classes located on lower Vancouver Island, British Columbia, Canada, was selected for this study. The subjects ranged in age from five years nine months to six years nine months at the beginning of the study.

Predictor Variables

Cognitive Development

Cognitive development involves a structured set of logico-mathematical abilities, processes and operations used to solve specific problem tasks. Special consideration being given to Arlin's (1981) observation that many previous studies gave weak justifications for their selections of cognitive development variables, six Piagetian-type tasks were selected for this study on the basis that the cognitive skills embodied in the tasks appear related to reading and mathematics and that the acquisition of these skills was a reasonable expectation of Grade One children. The specific tasks selected were Conservation of Number, Conservation of Solid Amount, Conservation of Liquid Amount, Conservation of Length, Seriation of Length and Horizontal Reference Frames (Yore & Ollila, 1984). Conservation refers to an individual's ability to recognize and rationalize the invariance of an identified quantity while other characteristics are changed. The logical thought embedded in conservation, specifically logical necessity, compensation, and reversibility, appear to reflect the logic needed to learn reading skills. Seriation, a logical grouping task, relates to the child's ability to identify and apply an ordered length relationship to a set of rods. Seriation appears to parallel the temporal sequence necessary for sentence and paragraph comprehension. Horizontal Reference Frame, an infralogical grouping task, deals with a child's conception of Euclidean space in which a consistent horizontal reference is viewed with changing, unrelated reference objects. The horizontal/vertical reference frame involves similar orientational discrimination as involved in the discrimination of letters, specifically; b, p, q, and d.

The tasks were administered using standard individual interview protocols and scored according to established criteria by a trained interviewer during September of Grade One. The scoring procedures utilized four ordered categories for the Conservation and Seriation tasks and five ordered categories for the Horizontal Reference Frame task.

Canadian Readiness Test

The CRT consists of seven subscales:

1. Technical Language of Literacy. Sixteen items designed to measure the child's knowledge of technical linguistic terms such as "letter," "word," etc. For example, the tester says, "Find each thing which is a 'letter'." The child draws a circle around all the "letters" found in an indicated row of symbols.
2. Word Matching. Twenty-one items in which the child selects from a display of four words the one which "looks just like" another word printed separately and specified by the tester.
3. Beginning Sounds. Nineteen items in which the child circles a picture whose name begins with the same initial phoneme as another word printed separately and specified by the tester.
4. Semantics. Twelve items measuring the child's ability to categorize. The task involves selecting from five pictures three which are conceptually identical.
5. Visual Letter Recognition/Letter-Name Knowledge. Twenty-one items measuring the child's ability to discriminate and recognize visually upper and lower case letters in relation to their letter-names. The tester names a letter and the child selects one of four alternatives.

6. Noun (learning rate on concrete words). Eighteen items to test the child's ability to select one of three printed words named by the tester. Prior to the test nine printed concrete nouns of varying length and configuration have been taught.
7. Non-noun (learning rate on abstract words). Eighteen items to test the child's ability to select one of three printed words named by the tester. Prior to the test, nine printed abstract non-noun words of varying length and configuration have been taught.

The CRT was administered at the start of the Grade One year during the middle of September by each classroom teacher. Testing periods varied in length from 30 to 45 minutes. No more than three subscales were administered during any one testing session. Kuder-Richardson reliability coefficients of the seven subscales ranged from .78 to .95 with the single most highly reliable measure being the test of Visual Letter Recognition/Letter-Name Knowledge.

Correlates

Reading Achievement

Four parts of the Stanford Achievement Test (Madden, Gardner, Rudman, Karlsen & Merwin, 1972) were administered at the end of Grade One to measure reading achievement. The reading subscales used were the Vocabulary, Reading Part A (word reading), Reading Part B (reading and comprehension) and Word Study Skills. The Stanford Achievement Test was administered and scored during May of the Grade One year by the classroom teachers according to the procedures in the test manual.

Mathematics Achievement

A school district-developed mathematics test battery, designed to measure mathematics achievement, was administered at the same time as the reading achievement tests. The mathematics battery consisted of four subtests.

1. Basic Facts and Computation. Thirty addition and subtraction items.
2. Numeration. Thirty-four items involving counting, associating numeral names and numerals, place value concepts, missing addend addition, inequality symbols, and multiplication basic facts.
3. Problem Solving. Fifteen word problems, predominantly requiring addition or subtraction.
4. Mathematical Concepts. Fourteen questions relating to money, geometry, time and comparative vocabulary.

Kuder-Richardson internal reliability coefficients ranged from .80 to .87 for the subscales. Content validity of the battery was supported through the correspondence of the items with the Grade One mathematics curriculum used in the school district.

Results

Correlational Analyses

Table 1 shows the intercorrelation matrix for the variables. Strong correlations can be seen between Solid, Liquid, and Length Conservation tasks, and within the CRT subscales, the mathematics battery, and the reading achievement subtests.

TABLE 1

Correlations Between Piagetian, CRT, Reading Achievement and Mathematics Achievement Variables

Subscales	Solid Cons.	Number Cons.	Liquid Cons.	Length Cons.	Horiz.	Seria.	Tech. Lang.	Letters	Word Match	Begin. Sds.	Seman.	Non-noun	Noun	Vocab.	Word Read.	Read. Comp.	Word Study	Comput.	Numbers	Prob. Solv.	Concept
Solid Cons.	—		.713**	.434**																	
Number Cons.		—								.467**				.351*				.369*		.508*	
Liquid Cons.			—	.435**																	
Length Cons.				—																	
Horizon					—																
Seriation						—															
Tech Lang							—	.504**		.368*	.360*	.432**	.378*	.390**	.403**	.354*		.376*			
Letters								—	.364*	.461**		.621**	.516**		.425**			.443**	.445**	.460**	
Word Match									—	.510**		.478**	.277		.360*			.376*	.433**		
Begin Sds.										—		.544**	.412**	.508**	.533**	.510**	.480**	.430**	.460**	.446**	
Semantics											—		.298	.364*				.657**	.550**	.600**	.378*
Non-noun												—	.719**	.409**	.647**	.518**	.552**	.396**	.479**	.464**	.455**
Noun													—	.475**	.566**	.493**	.529**	.400**	.509**	.469**	.334*
Vocabulary														—	.651**	.555**	.555**	.375*	.524**	.475**	
Word Read															—	.768*	.753**	.523**	.590**	.631**	
Read Comp																—	.644**	.584**	.671**	.630**	
Word Study																	—	.594**	.664**	.630**	
Computation																		.544**	.591**	.507**	
Numbers																		—	.677**	.710**	.354*
Prob. Solv.																			—	.737**	.398**
Concepts																				—	.499**
																					—

Note:
$N=99$
*$p<.05$, multistage Bonferroni test
**$p<.01$, multistage Bonferroni test

A series of canonical correlations was conducted in order to test the strength of the relationships between various sets of the variables. When the CRT subscales were taken as one set of predictors and the mathematics achievement tests were taken as the second set, a significant interrelationship was found ($F=3.849$, $df=28,250$, $p<.001$). A similar result occurred when the reading achievement tests were the correlates instead of the mathematics tests. Again, a significant relationship was found between the CRT and the reading achievement variables ($F=2.609$, $df=35,288$, $p<.001$).

In order to examine the impact of the addition of the Piagetian tasks to the predictive power of the CRT, a canonical correlation was conducted using the combined set of CRT and Piagetian scores as one group of predictors, and the mathematics achievement tests as the second group. A significant overall relationship was again found ($F=2.75$, $df=52,246$, $p<.001$). Because of this significant overall correlation an examination of the contribution of each of the subscales was justified and a series of univariate F-tests with (4,75) df was conducted. Each of the seven CRT subscales was significantly involved in the overall correlation ($p<.001$) but only two of the Piagetian tasks were significantly involved—Number Conservation ($p<.001$) and Seriation ($p<.01$).

A significant overall correlation was also found between the CRT-Piaget scores as one set of predictors and reading achievement scores as the second set ($F=1.85$, $df=65,297$, $p<.001$). Again because of this overall significance, an inspection of each of the series of univariate F-tests was justified. As before, all seven of the CRT subscales were significantly involved as predictors of reading achievement (Word Matching, $p=.015$; Semantics, $p=.03$; remaining subscales, $p<.01$) but only two Piagetian tasks were significantly involved in the association with reading achievement. These tasks were Number Conservation ($p<.01$) and Liquid Conservation ($p<.05$).

In summary, various canonical correlations showed that the seven reading readiness variables were consistently important in the interpretation of the canonical variables. Among the Piagetian tasks, Conservation of Number was the variable most often involved in the interpretation.

Factor Analysis

A final investigation involved the relationship between the underlying theoretical constructs which the CRT and the set of six Piagetian tasks were meant to represent. A principal components factor analysis of the intercorrelations between these two sets of measures was utilized to explore the pattern of underlying relationships. Table 2 shows the factor matrix after varimax rotation. It can be seen that the first factor, which accounts for 30% of the variance, is defined by the CRT. None of the Piagetian tasks is involved in this factor. Three of the Conservation tasks (Solid, Liquid, and Length) define the second factor, which accounts for 15% of the variance. None of the CRT variables is involved with this cluster. The third factor, accounting for 10.6% of the variance, is defined by Number Conservation, Seriation, Word Matching, Beginning Sounds, and Semantics. The fourth factor, accounting for 8% of the variance, is defined by a single variable—Horizontality.

Discussion

The results support the predictive validity of the CRT for both subsequent reading achievement and mathematics achievement.

TABLE 2

CRT and Selected Piagetian Tasks:
Results of Principal Components Analysis after Varimax Rotation

Variables	Factor 1	Factor 2	Factor 3	Factor 4
Solid Cons.		.851		
Number Cons.			.557	
Liquid Cons.		.850		
Length Cons.		.728		
Horizon				.916
Seriation			.657	
Tech. Lang.	.542			
Letters	.822			
Word Match			.541	
Begin. Sds.	.526		.585	
Semantics			.588	
Non-noun	.872			
Noun	.830			
Eigenvalues	3.94	1.99	1.38	1.04

Within the CRT battery, four of the seven subscales were the most powerful predictors. These four are Beginning Sound, Semantics, Noun, and Non-Noun. In the intial validation study of the CRT, (Evanechko, et al., 1973), the Letter Recognition subscale was the best single predictor of reading achievement, followed by Noun, Semantics, and Beginning Sounds. The letter Recognition subscale no longer serves as a significant predictor. A distinct ceiling effect was noted for this subscale ($X=18.23$, $S.D.=3.4$), perhaps indicating many children are now thoroughly exposed to letters before they start Grade One. Beginning Sounds, Semantics, and Noun learning rate remain as useful predictors of reading achievement.

The six selected Piagetian tasks, as indices of cognitive development, appear to be measuring three groups of competencies as indicated by the factor analysis, but contribute relatively little to the prediction of either reading or mathematics achievement when used in conjunction with the CRT. It is especially interesting that the CRT provides a better prediction of mathematics achievement than do the Piagetian tasks, given the strong theoretical rationale between the cognitive development paradigm and the child's development of mathematical concepts. These findings parallel Nelson's (1980) observations and have implications for mathematics education research and methodologies typically prescribed for early childhood mathematics, as both are frequently established around a cognitive development framework.

The relationship between the constructs underlying the CRT and those underlying the set of six Piagetian measures is interesting. It would seem that the clear separation between the CRT variables and four of the six Piagetian tasks would mean those four tasks would add extra information to the overall association between the predictor variables and the post-Grade One scores. However, the other two Piagetian tasks, Number Conservation and Seriation, are the ones which were significantly involved in the overall correlations between predictors and

achievement. Although the four Piagetian tasks (Solid, Liquid, and Length Conservation, and Horizontality) were chosen as providing a reasonable theoretical fit for cognitive skills needed in Grade One reading and mathematics, they were seen to measure something different from that which was measured by the CRT. Was the rationale for the selection of the specific Piagetian tasks faulty? As well as describing something different from the skills measured by the CRT, the dimensions described by the four Piagetian tasks do not seem to be measured by post-Grade One achievement tests. What are the implications of this lack of congruence between what is seen to be important in the Grade One curriculum, with respect to year-end testing, and what is involved cognitively in the various Piagetian tasks? As these tasks seem to be tapping something that is not reflected in post-Grade One reading and mathematics achievement scores, does this imply that we are neglecting to evaluate some important aspects of children's cognitive growth within our typical school achievement batteries? Or, alternatively, might these aspects of cognitive development be relatively unimportant in the actual learning of mathematics and reading?

One of the dimensions identified in the factor analysis, Factor 3, is a complicated factor which needs closer inspection. Variables with interpretable loadings on this factor include ones from both batteries which seemed most involved in the overall prediction of reading and mathematics achievement—Number Conservation and Seriation from the Piagetian set and Beginning Sounds and Semantics from the CRT. What is it that these measures have in common and why are they especially important with respect to subsequent school achievement? Subsequent study is needed to see if this result can be replicated but more importantly to see if it can be integrated into a theoretical framework which describes how children learn.

Taken together, the CRT and the set of six Piagetian tasks do form an impressive prediction of subsequent achievement, accounting for 56% of the variance in reading achievement and 65% of the variance in mathematics achievement. However, the Piagetian tasks contribute relatively little to this overall prediction, and add considerable logistic demands involved in the administration of the individual interviews. The CRT tests, used without the Piagetian measures, appear well supported as reliable and valid predictors of Grade One achievement in both reading and mathematics.

Note

1. Inquiries should be sent to Dr. Betty Collis, Department of Psychological Foundations, Faculty of Education, University of Victoria, P.O. Box 1700, Victoria, B.C., Canada, V8W 2Y2.

References

Arlin, P.K. (1981). Piagetian tasks as predictors of reading and mathematics readiness in Grades K-1. *Journal of Educational Psychology, 73* (5), 712-721.

Barrett, T.C. (1965). Visual discrimination tasks as predictors of first grade reading achievement. *The Reading Teacher, 18*, 276-282.

Bond, G.L., & Dykstra, R. (1967). The cooperative research program in first grade reading achievement. *Reading Research Quarterly, 2*, 5-141.

Brekke, B.W., & Williams, J.D. (1975). Conservation as a predictor of reading achievement. *Perceptual and Motor Skills, 40* (1), 95-98 .

Clymer, T., & Barrett, T.C. (1967). *Clymer-Barrett Prereading Battery.* Princeton, NJ: Personnel Press.

Dimitrovsky, L., & Almy, M. (1975). Early conservation as a predictor of later reading. *Journal of Psychology, 90,* 11-18.

Evanechko, P., Ollila, L., Downing, J., & Braun, C. (1973). An investigation of the reading readiness domain. *Research in the Teaching of English, 7* (1), 61-78.

Gates, A.I., Bond, G.L., & Russell, D.H. (1939). *Methods of determining reading readiness.* New York: Teacher College, Columbia University.

Hildreth, G.H., Griffiths, N.E., & McGauvran, M.E. (1965). *Metropolitan readiness test.* New York: Harcourt Brace Jovanovich.

Johnson, R.E. (1969). The validity of the Clymer-Barrett prereading battery. *The Reading Teacher, 22,* 609-614.

Jordan, V.B., & Brownlee, L. (1981, April). *Meta-analysis of the relationship between Piagetian and school achievement tests.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, CA.

Kingma, J., & Koops, W. (1983). Traditional intelligence and achievement tests. *British Journal of Educational Psychology, 53* (3), 178-290.

Madden, R., Gardner, E.F., Rudman, H.C., Karlsen, B., & Merwin, J.C. (1972). *Stanford achievement test.* New York: Harcourt Brace Jovanovich.

Nelson, D. (1980). *Studying problem solving behavior in early childhood.* (Lecture Series Award Monograph). Edmonton, Alberta: Faculty of Education, University of Alberta.

Smith, L.E. (1957). *Auditory discrimination and first grade reading achievement,* Unpublished Master's of Education Thesis, Illinois State Normal University.

Taloumis, T. (1979). Scores on Piagetian area tasks as predictors of achievement in mathematics over a four-year period. *Journal for Research in Mathematics Education, 10* (2), 120-134.

Waller, T.G. (1977). *Think first, read later: Piagetian prerequisites for reading.* Newark, DE: International Reading Association.

Yore, L.D., & Ollila, L.O., (1984). Cognitive growth and test-retest effects of selected conversation abilities, logical groupings and infralogical groupings. *Psychological Reports, 55,* 527-538.

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 12-36*

HORST H. MUELLER

ROBERT F. MULCAHY

LORRAINE WILGOSH

University of Alberta

BRONWYN WATTERS

Government of Northwest Territories

and

GABRIEL J. MANCINI

Edmonton Public School Board

An Analysis of WISC-R Item Responses With Canadian Inuit Children[1]

Item responses on nine WISC-R subtests (Object Assembly, Coding, Mazes excluded) from 273 Inuit children aged 7-0 to 8-11 years, 10-0 to 11-11 years, and 13-0 to 14-11 years in the Kitikmeot and Keewatin districts of Canada's Northwest Territories were examined with respect to item-total correlations and relative item difficulty. Overall, the WISC-R was found to be a reasonably internally consistent measure of the Wechsler intelligence constructs for these children. The majority of test items demonstrated significant correlations with the full test score, and also showed adequate and appropriate discrimination between the Wechsler verbal and performance constructs. While there was little evidence to suggest a need for major item revisions or reordering, the finding that fully one-third of the test items fell within the extreme deciles of relative difficulty for these children and, therefore, failed to add significantly to test variance is of some concern. The Verbal Scale tended to contain more items that were too difficult, whereas the Performance Scale tended to contain more items that were too easy. The implications of these findings are discussed with respect to the WISC-R's use as a global intelligence measure with Inuit children.

Over the last quarter-century, the Wechsler Intelligence Scale for Children (WISC) and its 1974 revision, the WISC-R, have found increasing acceptance with Canadian psychologists. Today the WISC-R is clearly the "test of choice" for the assessment of the intellectual functioning of individual Canadian children. However, along with this increased popularity has also come increasing concern with re-

---

Mr. Mueller is presently a doctoral student in the Department of Educational Psychology, University of Alberta, and a Research Associate with Western Industrial Research and Training Centre, Edmonton.

Dr. Mulcahy and Dr. Wilgosh are Professors in the Department of Educational Psychology at the University of Alberta.

Ms. Watters is the Chief of Social Program Services, Department of Social Services, Government of the Northwest Territories, Yellowknife.

Dr. Mancini is a Consultant, Behavior Disorders, with the Edmonton Public School Board.

gard to whether using an American-normed psychometric instrument results in significant cultural bias against Canadian majority or ethnic and racial minority children.

A number of studies have examined the effects of "Canadianizing" WISC Information items (Spreen & Tryk, 1970), and WISC-R Information and Comprehension items (Vernon, 1974, 1976, 1977). Recently, Marx (1984) compared the item performance of two samples of Canadian children administered Vernon's (1977) modified WISC-R Information subtest to item difficulty data obtained for Information from the 1974 WISC-R standardization. Marx concluded that "Vernon's suggested changes to the WISC-R Information subtest resulted in higher raw scores for this subtest for both clinical and nonclinical samples" (Marx, 1984, p. 34). Such results are suggestive of some WISC-R cultural bias. But, as yet, there is no research comparing WISC-R item performance of a normative sample of Canadian children to the WISC-R standardization data using a standard administration, nor is there research directly comparing performance on the "Canadianized" WISC-R to the standard WISC-R in a normative Canadian sample.

The majority of studies examining the standard WISC-R performance of Canadian majority children have focused on more global subtest and Scale score patterns of differential responding rather than on the analysis of individual item responses (e.g., Dash, Dennis, Mueller, Mancini, Snart, & Short, 1983; Peters, 1976; Wersh & Briere, 1981). Such studies have failed to clearly establish WISC-R performance differences attributable to cultural differences between Canadian normal and clinic-referred children, and the American standardization sample (see Dash et al., 1983). While Canadian majority children may demonstrate some increased difficulty with a few specific WISC-R items, their overall performance does not appear significantly affected.

But what of Canada's ethnic and racial minority children? Might not the WISC-R and its norms be significantly biased against these groups? Although the relatively abundant American research literature on WISC-R bias with American minorities (e.g., Mishra, 1982; Sandoval, 1979; Sandoval & Miille, 1980; Sandoval, Zimmerman, & Woo-Sam, 1983) has generally found little or no overall test bias, one cannot simply generalize such findings to the Canadian situation, where the research has focused almost exclusively on the indigenous Native population.

Although rather sparse and fragmentary, research and ancedotal data support the generalization that Canadian Native children's performance on tasks assessing visual-perceptual and nonverbal analytic skills compares quite favorably to that of Canadian majority children, but compares poorly on tasks that assess verbal-conceptual and linguistic skills (e.g., Kleinfield, 1973; MacArthur, 1968; St. John, Krichev, & Bauman, 1976; Taylor & Skanes, 1976).

Seyfort, Spreen, and Lahmer (1980) evaluated WISC-R item responses and difficulty gradients in a sample of 177 Native children aged 6 to 13 years living on reserves and attending regular schools in Southwestern British Columbia. Single item response frequency and difficulty gradients were analyzed by coding items as 0 for incorrect and 1 for correct responses; partial or qualitative scores were not recorded separately. Their data confirmed previous findings insofar as their sample displayed average visual-perceptual analysis skills and relatively poor verbal skills. But of more significance, Seyfort et al. found that only approximately 60% of the 157 test items answered correctly by more than one child correlated significantly ($p<.05$) with own subtest total score. This apparent lack of "validity" (p. 18), as indexed by test-total correlations for many of the WISC-R items, suggests that either these Native children were having some difficulty in properly understanding

many of the items, or that many WISC-R items were tapping rather different skills or abilities than in the standardization sample. Only approximately 25% of the items answered correctly by at least one child were of moderate difficulty, whereas approximately 43% of such scorable items fell within the extreme deciles of relative difficulty. However, any analysis of WISC-R item difficulty that collapses data across a broad age range will find a rather large number of items to have a mean difficulty index outside the "desirable" range of .30 to .70 suggested by Anastasi (1976) because no single item is designed to be equally difficult to children for all ages. A statement regarding inappropriate difficulty levels for specific WISC-R items requires item response data from less broad age groups.

The purpose of the present study was to examine WISC-R item-total correlations and relative item difficulty for three age pools of Canadian Inuit children drawn from a more complete norming sample. It was believed that such an examination of the item response and difficulty gradient of the WISC-R would facilitate more accurate and meaningful assessment of Inuit children in Canada's Northwest Territories.

Method

Subjects

The subjects used for the present analysis were all children ($N=273$; 137 male + 136 female) aged 7-0 to 8-11 years, 10-0 to 11-11 years, and 13-0 to 14-11 years included within a larger sample of children assessed as part of a psychometric test norming project in Canada's Northwest Territories in 1981 (see Mulcahy & Watters, 1983[2]). The original normative sample consisted of 366 students (185 males, 181 females) from 13 villages in the Kitikmeot and Keewatin districts of the Northwest Territories. These 366 children were randomly selected from the complete school registry after stratification by age and sex. The population of these regions is approximately 90% Inuit, 8.5% Caucasian, and 1.5% Dene. This normative sample, the sampling procedures used and the testing procedures are described in detail by Mulcahy & Watters.

Procedures

These data were collected between October 15 and November 15, 1981. The WISC-R was administered by nine trained psychometricians (4 male, 5 female). All testers received extra training with respect to the assessment of northern Native children prior to travelling to the Territories. All subjects were individually tested in their school location with some testing occurring on non-school days or early evenings.

A total of seven WISC-R items in the Information (5), Similarities (1), and Comprehension (1) subtests were modified to better reflect Canadian content (see Vernon, 1974, 1976, 1977). Also, in consultation with a practicing psychologist in the Northwest Territories with extensive experience in administering the WISC-R to Native children, it was decided to modify slightly the wording of three further items (one in Similarities and two in Arithmetic) to enhance understanding. Minor modifications to the query procedures outlined in the *WISC-R Manual* (Wechsler, 1974) were also made for the Similarities, Vocabulary, and Picture Completion subtests.[3] In all other respects, the WISC-R was administered in the standard manner.

Scoring Reliability. Testers were paired and each checked the other's WISC-R protocols for correct item scoring and summation. As well, every protocol was further checked by the second author for correct scoring. Since receptive and expressive language deficiency is known to interfere with WISC-R performance, an at-

tempt was made to test only those selected children who, according to teachers and testers, were "sufficiently proficient" in the English language. All testers were also instructed to record on each test protocol any concerns that a lack of receptive or expressive language skill might have interfered with adequate performance. A total of 32 such protocols were determined invalid upon closer scrutiny by the authors and were deleted from the norming sample. Thus a total of 366 WISC-R protocols were retained for norming analysis. For the present item analyses, the protocols of all children aged 7-0 to 8-11 years were pooled into one group (*G1;* $n=71$), all children aged 10-0 to 11-11 years were pooled into another group (*G2;* $n=101$), and all children aged 13-0 to 14-11 were pooled into a third group (*G3;* $n=96$).

Data Analysis

Since ANOVAs on subtest total scores by age and sex carried out on the original norming sample (see Mulcahy & Watters[2]) revealed no significant age by sex interactions or sex main effects, it was decided to collapse data across sex for all item analyses.

Single item response frequencies and item difficulty gradients were generated by coding all item responses as either 0 for incorrect or 1 for correct. Items were considered correctly answered if any credit could be allowed. While this procedure functions to reduce response variance and lower item-total correlations, it was followed to increase the ordinal consistency of test items and to simplify data analysis (Seyfort et al., 1980). No error is created by such a scoring procedure for the Information, Picture Completion, and Arithmetic subtests, and only a relatively small amount of error is created for the Similarities, Vocabulary, Comprehension, and Digit Span subtests. Significantly greater error is introduced for such subtests as Picture Arrangement, Block Design, Object Assembly and Mazes wherein children can obtain scores greater than two for items. But for Picture Arrangement and Block Design only variance due to speed of performance is completely lost by using a "0/1" scoring procedure. It is only with Object Assembly and Mazes that too much variance due to "correctness of performance" is lost to permit a reasonably accurate analysis of relative item difficulties. Therefore, these latter two subtests were excluded from the present analysis. The Coding subtest was also excluded from analysis because it does not contain an increasing item difficulty component. Item analysis, therefore, was carried out on only the 191 items within the remaining six Verbal subtests (142 items) and three Performance subtests (49 items).

A count of the number of children in each age pool correctly answering each test item and a calculation of the relative proportion of each pool correctly answering each test item were performed. Test items in each subtest were ranked in order of relative difficulty on the basis of the proportion of children in each age pool who answered the item correctly. These data were used for the item difficulty analyses. The Spearman *rho* formula was used to calculate the rank-order correlation between the standardized subtest item order and a revised order based on obtained difficulty data for each subtest at each of the three age levels. The N used for each of these calculations was the total number of subtest items answered by at least one child within each age pool. Such items will be termed "scorable items" for the purposes of this paper.

Item scores (0/1) were also used to generate a set of three Product-Moment correlations (PM) for each of the 191 items being analyzed. Each test item was correlated with a "Verbal Score" consisting of the sum of Information, Similarities, Arithmetic, Vocabulary and Comprehension subtest scores; with a "Performance Score" consisting of the sum of Picture Completion, Picture Arrangement, and

Block Design subtest scores; and a "Full Score" consisting of the sum of the Verbal and Performance Scores. These item-total correlation coefficients were interpreted as discrimination indices related to test construct validity at the item, rather than subtest, level (see Vance, Gaynor, & Coleman, 1977; Seyfort et al., 1980).

Results and Discussion

Subtest Item Difficulty and Validity

Tables 1 through 4 present the results of the item analysis on each of nine WISC-R subtests for each of three age pools. These results have been tabulated at each age level into six columns as follows: 1) number of subjects who correctly answered the item *(f)*, 2) the relative proportion of subjects who correctly answered the item *(P)*, 3) an obtained item order based on relative item difficulty *(R)*, 4) the PM correlation between the subtest item and the Verbal Score (*r*VS), 5) the PM correlation between the item and the Performance Score (*r*PS), and 6) the PM correlation between the item and the Full Score (*r*FS). Items are tabulated in the same sequence as given within the *WISC-R Manual* (Wechsler, 1974). Spearman rank-order correlations between standardized item order and revised orders based on obtained data are also presented at the bottom of each table. For reasons of parsimony, only scorable items in any age pool are tabulated, as can be determined with reference to the tables.

Two aspects of test items will be discussed for each of the nine subtests at each of the three age levels: item validity and item relative difficulty. Relative item difficulty is indexed by the proportion of subjects taking a test who correctly answer the item. Items correctly answered by between 30 and 70 percent of subjects are considered moderately difficult and desirable from a test construction point of view. Correctly answered items falling within the extreme deciles of difficulty ($P=0$ to .10, or .90 to 1.00) are either too easy or difficult for the majority of children and fail to add significantly to total test variance (see Anastasi, 1976). Item validity may be inferred from the degree of relationship between item responses and some criterion of performance within or external to the test (see Anastasi, 1976). For the present analysis, a traditional item difficulty index was generated by determining the relative proportion of children in each age pool who answered an item correctly, and three separate indices of item validity were generated by correlating responses to each item with Verbal score, Performance Score, and Full Score.

Overview

For ages 7-0 to 8-11 (G1), of the 191 items administered, only 130 (83 Verbal, 47 Performance) were scorable. The corresponding number for ages 10-0 to 11-11 (G2) are 155 items (106 Verbal, 49 Performance), and 13-0 to 14-11 (G3), 170 items (121 Verbal, 49 Performance). For G1, only 24% (G2: 24%; G3: 20%) of Verbal items and 43% (G2: 24%; G3: 10%) of Performance items are of moderate difficulty (P=.30 to .70). For G1, for the Verbal Scale, as many as 34% (G2: 31%; G3: 31%) of scorable items fall within the lower decile of difficulty (too hard), whereas only 5% (G2: 13%; G3: 21%) fall within the upper decile (too easy). In contrast, only 15% (G2: 8%; G3: 2%) of Performance Scale items fall within the lower decile and 17% (G2: 31%; G3: 49%) fall within the upper decile. These items failed to add significantly to the total test variance for these Inuit children.

For all three groups, the majority of scorable test items demonstrate a pattern of item-total correlations consistent with the WISC-R's construction. For the Verbal Scale, for G1, 72% (G2: 69%; G3: 69%) of scorable subtest items correlate significantly with the Verbal Score while only 34% (G2: 31%; G3: 25%) correlate significantly with the Performance Score. Similarly, 70% (G2: 73%; G3: 65%) of scorable Performance Scale items correlate significantly with the Performance

Score and only 30% (G2: 31%; G3: 22%) correlate significantly with the Verbal Score. A total of 68% of all 130 (G2: 67% of 155; G3: 59% of 170) scorable test items correlate significantly with the Full Score.

Information Subtest. Examination of Table 1a for the 7-8 year pool (G1) reveals that 73% (G2: 81%; G3: 79%) of scorable items correlate significantly with the Verbal Score. Only two items (G2: 2 items; G3: 3 items) also correlate significantly with the Performance Score. This pattern suggests that Information is a reasonably good measure of the WISC-R verbal construct for each age group. Table 1a also reveals that Information is only moderately well ordered at age 7-8 with respect to obtained item difficulty (*rho*=.84; G2: *rho*=.94; G3: *rho*=.96), although only one item, #8 (*"Name . . . after March."*), is misplaced by more than two ranks from its standard placement (easier).[4] For G2, four scorable items are misplaced by more than two rank-positions (all easier); item #8, #24 (*"How tall . . . Canadian man?"*), and #19 (*"What are the names . . . oceans . . . Canada?"*), modified to be more Canadian in content, and #25 (*"What . . . barometer?"*). For G3, only three scorable items are misplaced by more than two ranks, item #19 (easier), #12 (*"Who . . . America?"*), an unchanged American-content item (harder), and #20 (*"How many . . . ton?"*; harder). One scorable Information item is of moderate difficulty for G1 (G2: 29%; G3: 13%) while 55% (G2: 62%; G3: 63%) fall within the extreme deciles of difficulty. A sharp increase in item difficulty is apparent for G1 beginning with item #4 (G2: #10; G3: #12).

Similarities Subtest: Table 1b reveals that, for G1, 83% (G2: 92%; G3: 88%) of the scorable Similarities items correlate significantly with the Verbal Score. However, fully 67% (G2: 69%; G3: 47%) of scorable items also show significant positive correlation with the Performance Score and, therefore, may be considered somewhat weak in scale discrimination. At 7-8, there is reasonably good consistency in obtained item placement compared to the standardized item order (*rho*=.93); with only one scorable item misplaced by more than two ranks from standard (easier), #15 (*"first-last"*). The obtained item order for G2 is only slightly discrepant from the standardized order (*rho*=.89; G3: *rho*=.94); with two scorable items (G3: only item #3) misplaced by more than two ranks (harder), #3 (*"shirt-hat"*) and #6 (*"beer-wine"*). For G1, one scorable item is of moderate difficulty (G2: 62%; G3: 41%), while fully 75% (G2: 31%; G3 30%) of scorable items fall within the lower decile of difficulty (too hard). No items fall within the upper decile (too easy) for any age group. For all 3 groups, a dramatic increase in item difficulty is apparent starting with item #10 (*"pound-yard"*).

Arithmetic Subtest. Table 2a shows that, for G1, 71% (G2: 56%; G3: 38%) of the scorable arithmetic items correlate significantly with the Verbal Score, but 43% (G2: 50%; G3: 19%) of Arithmetic items also correlate significantly with the Performance Score. Thus Arithmetic appears to be a rather poor discriminator between the WISC-R verbal and nonverbal constructs, particularly at age 10-11, and at 13-14, arithmetic does not appear to be a particularly strong measure of any of the three WISC-R scales or constructs. Arithmetic items are quite well ordered with respect to obtained relative difficulty for all groups (G1: *rho*=.98; G2: *rho*=.997; G3: *rho*=.97) with no items misplaced by more than two ranks from standard for G1 and 2, and only item #6 significantly misplaced from standard for G3. For G1, arithmetic contains a relative large proportion of scorable items (38%) (G2: 1 item; G3: 3 items) of moderate difficulty, but another 38% (G2: 67%; G3: 50%) of scorable items also fall within the extreme deciles of difficulty and fail to add greatly to total subtest variance. A moderately large increase in difficulty occurs for G1 with item #10 (G2: #15).

TABLE 1

Frequency and Relative Probability of Correct Response, Item Order Based on Probability of Correct Response, and Correlation of Items With Verbal Score, Performance Score, and Full Score for Three Age Pools

(a) Information Subtest

	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
Item	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS
1	54	.76	3	.45	.34	.45	97	.96	2	**.09**	**.03**	**.06**	96	1.00	1	n/c	n/c	n/c
2	68	.96	1	**.01**	**.22**	**.13**	101	1.00	1	n/c	n/c	n/c	96	1.00	2	n/c	n/c	n/c
3	64	.90	2	.40	**.27**	.37	97	.96	3	**.01**	**.03**	**.02**	96	1.00	3	n/c	n/c	n/c
4	19	.27	5	.49	**.18**	.37	53	.53	5	.54	**.13**	.38	87	.91	4	.37	.35	.42
5	11	.16	7	.32	**.08**	**.21**	52	.52	6	.49	**.23**	.41	85	.89	5	.42	**.18**	.35
6	13	.18	6	.57	**.25**	.45	38	.38	8	.58	**.20**	.44	81	.84	6	.42	**.18**	.35
7	9	.13	8	.50	.38	.49	46	.46	7	.54	.31	.49	77	.80	8	.41	.28	.40
8	27	.38	4	.44	**.15**	.32	69	.68	4	.42	**.07**	.28	79	.82	7	.29	**.09**	**.22**
9	6	.09	9	.40	**.18**	.32	33	.33	9	.43	**.09**	.30	48	.50	9	.38	**.14**	.31
10	1	.01	10	**.26**	**.16**	**.23**	13	.13	10	.52	*.21*	.42	41	.43	10	.60	.30	.53
11	0	0	12	n/c	n/c	n/c	13	.13	11	.44	.37	.47	34	.35	11	.52	**.21**	.43
12	0	0	13	n/c	n/c	n/c	4	.04	13	.51	**.10**	.35	7	.07	15	.33	**.22**	.32
13	1	.01	11	**.08**	**-.02**	**.03**	5	.05	12	.49	**.22**	.41	11	.12	12	.48	**.19**	.39
14	0	0	14	n/c	n/c	n/c	4	.04	14	.46	**.08**	.30	11	.12	13	**.24**	**.09**	**.19**
15	0	0	15	n/c	n/c	n/c	2	.02	.15	**.15**	**.18**	**.19**	7	.07	16	.53	**.21**	.43

TABLE 1 continued

Item	AGE 7 & 8 YEARS (*N*=71)						AGE 10 & 11 YEARS (*N*=101)						AGE 13 & 14 YEARS (*N*=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
16	0	0	16	n/c	n/c	n/c	2	.02	16	.31	**.06**	**.21**	5	.05	17	.41	**.10**	.30
17	0	0	17	n/c	n/c	n/c	1	.01	18	.24	**.22**	.27	3	.03	18	.36	**.03**	**.23**
18	0	0	18	n/c	n/c	n/c	1	.01	19	.34	**-.02**	**.18**	2	.02	19	**.10**	**.05**	**.09**
19	0	0	19	n/c	n/c	n/c	2	.02	17	.41	**.15**	.32	9	.09	14	.37	**.16**	.31
20	0	0	20	n/c	n/c	n/c	0	0	22	n/c	n/c	n/c	1	.01	23	.28	**.14**	**.24**
21	0	0	21	n/c	n/c	n/c	0	0	23	n/c	n/c	n/c	2	.02	20	.29	**.10**	**.23**
22	0	0	22	n/c	n/c	n/c	0	0	24	n/c	n/c	n/c	2	.02	21	.29	**.01**	**.18**
23	0	0	23	n/c	n/c	n/c	0	0	25	n/c	n/c	n/c	1	.01	24	.28	**.14**	**.24**
24	0	0	24	n/c	n/c	n/c	1	.01	20	.34	**-.02**	**.18**	2	.02	22	.26	**.16**	.25
25	0	0	25	n/c	n/c	n/c	1	.01	21	.34	**-.02**	**.18**	0	0	25	n/c	n/c	n/c
Rho:			.84						.94						.96			

TABLE 1 continued

(b) Similarities Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
1	30	.42	1	.54	.35	.50	76	.75	1	.50	.26	.44	82	.85	1	.29	**.22**	.29
2	20	.28	2	.61	.39	.56	51	.51	4	.57	.25	.47	74	.77	2	.41	.27	.40
3	16	.23	4	.43	**.18**	.33	42	.42	6	.57	.35	.53	45	.47	8	.42	**.22**	.37
4	17	.24	3	.61	.31	.51	53	.53	2	.50	.38	.51	61	.64	4	.54	.42	.56
5	16	.23	5	.50	.42	.52	52	.52	3	.63	.45	.63	64	.67	3	.47	**.22**	.40
6	13	.18	6	.58	.35	.52	32	.32	9	.53	.31	.49	46	.48	7	.60	.32	.54
7	13	.18	7	.64	.33	.54	47	.47	5	.53	.36	.52	54	.56	5	.56	.26	.48
8	9	.13	8	.62	.33	.53	35	.35	7	.64	.31	.55	50	.52	6	.65	.43	.63
9	7	.10	9	.33	.35	.39	33	.33	8	.59	.36	.55	45	.47	9	.60	.37	.56
10	0	0	13	n/c	n/c	n/c	3	.03	11	.33	**.13**	.26	12	.13	12	.38	**.22**	.35
11	1	.01	11	**.26**	**.16**	**.23**	3	.03	12	.35	**.04**	**.22**	15	.16	10	.60	.35	.56
12	2	.03	10	.38	**.10**	**.26**	5	.05	10	.43	**.09**	.30	15	.16	11	.49	**.08**	.34
13	0	0	14	n/c	n/c	n/c	1	.01	13	**.10**	**.11**	**.12**	9	.09	13	.47	**.15**	.37
14	0	0	15	n/c	n/c	n/c	0	0	14	n/c	n/c	n/c	2	.02	15	.32	**.19**	.30
15	1	.01	12	**.25**	**-.01**	**.13**	0	0	15	n/c	n/c	n/c	8	.08	14	.45	.34	.45
16	0	0	16	n/c	n/c	n/c	0	0	16	n/c	n/c	n/c	1	.01	16	**.23**	**.18**	**.24**
17	0	0	17	n/c	n/c	n/c	0	0	17	n/c	n/c	n/c	1	.01	17	**.23**	**.18**	**.24**
Rho:			.93						.89						.94			

For Tables 1 to 4:

1. Correlation coefficients that are not calculable are marked n/c.
2. Correlations printed in **bold** are **non**significant at the .01 level.
3. Rho is the rank-order between the standardized item order and a revised item order based on obtained item difficulty indices.

TABLE 2

Frequency and Relative Probability of Correct Response, Item Order Based on Probability of Correct Response, and Correlation of Items With Verbal Score, Performance Score, and Full Score for Three Age Pools

(a) Arithmetic Subtest

	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
Item	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS
1	71	1.00	1	n/c	n/c	n/c	101	1.00	1	n/c	n/c	n/c	96	1.00	1	n/c	n/c	n/c
2	65	.92	2	.31	**.23**	.30	99	.98	2	**.22**	**.17**	**.23**	96	1.00	2	n/c	n/c	n/c
3	63	.89	3	.28	**.14**	**.23**	99	.98	3	**.12**	**-.01**	**.06**	96	1.00	3	n/c	n/c	n/c
4	42	.59	6	.35	.28	.36	92	.91	6	.32	.31	.37	95	.99	5	**.12**	**.23**	**.20**
5	44	.62	5	.33	**.26**	.33	93	.92	5	**.14**	**.08**	**.13**	94	.98	6	.25	**.19**	.25
6	52	.73	4	.50	.30	.45	96	.95	4	.27	.26	.31	96	1.00	4	n/c	n/c	n/c
7	41	.58	7	.54	.31	.47	80	.79	7	.41	.26	.38	94	.98	7	**.09**	.28	**.21**
8	28	.39	8	.59	.35	.52	77	.76	8	.32	.34	.38	89	.93	8	**.10**	**-.06**	**.02**
9	23	.32	9	.55	.34	.49	58	.57	9	.47	.31	.45	77	.80	9	.29	**.19**	.28
10	10	.14	10	.40	**.20**	.33	29	.29	10	.37	.35	.42	61	.64	10	.31	**.10**	**.24**
11	8	.11	11	.52	.37	.49	25	.25	11	.37	.36	.42	60	.63	11	.27	**.08**	**.21**
12	2	.03	12	**.09**	**.19**	**.16**	15	.15	12	.41	**.18**	.34	42	.44	12	.35	**.21**	.33
13	0	0	15	n/c	n/c	n/c	8	.08	13	.45	.26	.41	19	.20	14	.47	.28	.44
14	1	.01	14	**.03**	**.11**	**.08**	6	.06	14	.38	.31	.40	20	.21	13	**.17**	**.16**	**.19**
15	2	.03	13	**.16**	**.01**	**.09**	2	.02	15	**.17**	**.16**	**.19**	4	.04	16	**.08**	**.06**	**.08**
16	0	0	16	n/c	n/c	n/c	2	.02	16	**.13**	**.21**	**.20**	6	.06	15	**.19**	.30	.28
17	0	0	17	n/c	n/c	n/c	1	.01	17	**.04**	**.17**	**.13**	0	0	17	n/c	n/c	n/c
18	0	0	18	n/c	n/c	n/c	1	.01	18	**.04**	**.17**	**.13**	0	0	18	n/c	n/c	n/c
Rho:			.98						.99							.97		

TABLE 2 continued

(b) Vocabulary Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS
1	55	.78	1	.57	.31	.49	90	.89	1	.40	**.19**	.34	92	.96	1	.35	**.17**	.30
2	51	.72	3	.60	.32	.51	84	.83	4	.30	**.16**	.26	91	.95	2	.27	**.14**	**.24**
3	51	.72	4	.47	**.25**	.40	87	.86	2	.38	**.19**	.33	88	.92	4	.39	**.17**	.33
4	52	.73	2	.56	**.15**	.38	87	.86	3	.42	**.15**	.32	88	.92	5	.38	**.24**	.36
5	46	.65	6	.57	.41	.55	83	.82	5	.45	**.17**	.36	89	.93	3	.32	**.18**	.30
6	50	.70	5	.58	**.22**	.44	81	.80	6	.50	**.16**	.38	83	.87	6	.46	**.20**	.38
7	31	.44	7	.63	.30	.51	66	.65	7	.42	**.02**	.25	80	.83	7	.53	.36	.52
8	26	.37	8	.54	.34	.49	64	.63	8	.52	.37	.51	79	.82	8	.31	.32	.36
9	5	.07	12	.49	**.24**	.40	20	.20	13	.58	.31	.51	63	.66	10	.59	.46	.61
10	14	.20	9	.51	**.27**	.43	45	.45	9	.49	**.11**	.34	64	.67	9	.59	.31	.53
11	10	.14	11	.57	**.23**	.44	25	.25	11	.60	**.17**	.44	42	.44	13	.48	**.15**	.37
12	13	.18	10	.50	**.20**	.38	32	.32	10	.51	**.23**	.43	60	.63	11	.49	**.24**	.43
13	5	.07	13	.56	**.15**	.38	25	.25	12	.58	**.18**	.43	47	.49	12	.41	**.18**	.35
14	4	.06	14	.45	**.16**	.33	3	.03	15	.42	**.23**	.37	22	.23	14	.64	.32	.57
15	0	0	16	n/c	n/c	n/c	3	.03	16	.50	**.13**	.35	11	.12	15	.54	**.16**	.41
16	0	0	17	n/c	n/c	n/c	2	.02	18	.41	**.15**	.32	7	.07	18	.53	.33	.50
17	1	.01	15	**.25**	**-.01**	**.13**	5	.05	14	.50	**.14**	.37	7	.07	19	.34	**.16**	.30
18	0	0	18	n/c	n/c	n/c	0	0	19	n/c	n/c	n/c	1	.01	21	.28	**.14**	**.24**
19	0	0	19	n/c	n/c	n/c	0	0	20	n/c	n/c	n/c	8	.08	17	.48	**.16**	.38
20	0	0	20	n/c	n/c	n/c	2	.02	17	.41	**.15**	.32	10	.10	16	.54	**.17**	.42
21	0	0	21	n/c	n/c	n/c	0	0	21	n/c	n/c	n/c	0	0	24	n/c	n/c	n/c

TABLE 2 continued

Item	AGE 7 & 8 YEARS (*N*=71)						AGE 10 & 11 YEARS (*N*=101)						AGE 13 & 14 YEARS (*N*=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
22	0	0	22	n/c	n/c	n/c	0	0	22	n/c	n/c	n/c	0	0	25	n/c	n/c	n/c
23	0	0	23	n/c	n/c	n/c	0	0	23	n/c	n/c	n/c	3	.03	20	.25	**.01**	**.15**
24	0	0	24	n/c	n/c	n/c	0	0	24	n/c	n/c	n/c	1	.01	22	.28	**.24**	.30
25	0	0	25	n/c	n/c	n/c	0	0	25	n/c	n/c	n/c	1	.01	23	**.10**	**-.13**	**-.01**
Rho:			.95						.94						.97			

Vocabulary Subtest. Examination of Table 2b reveals that, for G1, 93% (G2: 100%; G3: 96%) of the scorable Vocabulary items correlate significantly with the Verbal Score, but 33% (G2: 2 items; G3: 26%) of scorable Vocabulary items also correlate significantly with the Performance Score. Item rankings based on obtained relative item difficulty are close to standardized placement for G1 (*rho*=.95); with only one item deviating from standard by more than two ranks (harder), #9 (*"thief"*). Vocabulary appears to be a very strong and efficient discriminator between verbal and nonverbal skills for age 10-11 Inuit children. Vocabulary items are also relatively well ordered for this group (*rho*=.95); with only two scorable items misplaced by more than two ranks from standard, #9 (harder) and #17 (*"nuisance"*; easier). For G3, *rho*=.96 and four scorable items are misplaced by more than two ranks, #20 (*"migrate"*), #23 (*"mantis"*), #25 (*"belfry"*), all easier, and #18 (*"harder"*; harder). For G1, only 27% (G2: 22%; G3: 22%) of scorable Vocabulary items are of moderate difficulty, and another 27% (G2: 28%; G3: 22%) fall within the lower decile of difficulty (too hard). No items fall within the upper decile for G1 and G2. A large increase in relative difficulty is apparent with item #9 for G1.

Comprehension Subtest. Table 3a shows that, for G1, all of the scorable Comprehension items (G2: 67%; G3: 65%) correlate significantly with the Verbal Score, and only two items (13%) also correlate significantly (G2: 18%; G3: 13%) with the Performance Score, giving the Comprehension subtest relatively good overall scale discrimination. Table 3a also reveals quite a bit of inconsistency in obtained item placement compared to the standardized placement (*rho*=.77) at 7-8; with six scorable items misplaced by more than two ranks from standard, #8 (*"Why . . . licence plates?"*), #10 (*"Why . . . on letters?"*), #11 (*"Why . . . meat packing plants?"*), all harder, #14 (*"How are . . . hardcover books?"*), #15 (*"Why . . . promise be kept?"*), and #16 (*"Why . . . in cloth?"*), all easier. Comprehension items are moderately well ordered with respect to difficulty for G2 (*rho* .87; G3: *rho*=.93); with four items (G3: only 2 items, #11 + #12) misplaced by more than two ranks from standard, #11 and #12 (*"Why . . . street beggar?"*), both harder, and #15 and #16 (both easier). Approximately 38% (G2: 2 items; G3: 2 items) of scorable Comprehension items are of moderate difficulty for G1 while another 56% (G2: 2 items) fall within the lower decile of difficulty (too difficult). No items fall within the upper decile of difficulty for G1 and G2. For G3, 29% of scorable items fall within the extreme deciles of difficulty. A moderate increase in item difficulty is apparent beginning with item #7 (*"In what ways . . . of wood?"*), for all three age pools.

Digit Span Subtest. Table 3b reveals that, for G1, only 50% (G2: 1 item; G3: 2 items) of the Digits Forward scorable items correlate significantly with the Verbal Score, and only one item (G2 and G3: 0 items) correlates significantly with the Performance Score, although the majority of these items (G2: only 2 items; G3: only 3 items) do demonstrate moderate positive correlation with the Full Score at age 7-8. These item-total correlate patterns suggest that Digits Forward is not a particularly good measure of any of the WISC-R constructs with the Inuit. Digits Forward items are quite well ordered with respect to obtained item difficulty (G1: *rho*=.93; G2: *rho*=.96; G3: *rho*=.98); with no items displaced by more than two ranks from standard. For G1, only one item is of moderate difficulty, whereas 38% of scorable items fall within the lower decile of difficulty (too hard). No items appear too easy. A dramatic increase in relative difficulty is apparent after the second item-pair (5 digits). For G2, only one item-pair (#5 & 6) is of moderate difficulty while the first two pairs (#1 & 2, 3 & 4) fall within the upper decile of difficulty (too easy) and the fifth item-pair (#9 & 10) falls within the lower decile (too

TABLE 3

Frequency and Relative Probability of Correct Response, Item Order Based on Probability of Correct Response, and Correlation of Items With Verbal Score, Performance Score, and Full Score for Three Age Pools

(a) Comprehension Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS
1	44	.62	1	.54	**.22**	.42	84	.83	1	.33	**.12**	.26	87	.91	2	.31	**.22**	.31
2	33	.47	3	.59	**.13**	.39	74	.73	4	.33	**.02**	**.19**	81	.84	3	.37	**.13**	.29
3	33	.47	4	.60	**.23**	.45	77	.76	3	.47	.25	.41	88	.92	1	**.24**	**.18**	.25
4	38	.53	2	.68	.34	.57	79	.78	2	.38	**.18**	.32	77	.80	5	.36	**.17**	.31
5	28	.39	5	.57	**.19**	.41	58	.57	6	.39	**-.01**	**.22**	68	.71	6	.35	**.30**	.38
6	24	.33	6	.46	**.25**	.40	59	.58	5	.54	**.20**	.42	79	.82	4	.48	**.19**	.39
7	7	.10	8	**.26**	**.10**	**.20**	25	.25	7	.45	**.22**	.38	33	.34	8	.59	**.21**	.47
8	4	.06	13	.51	**.21**	.39	13	.13	9	.40	.28	.39	29	.30	9	.57	**.15**	.43
9	12	.17	7	.51	.28	.44	23	.23	8	.54	.26	.46	50	.52	7	.59	.30	.52
10	4	.06	12	.29	**.17**	**.26**	12	.12	10	**.04**	**-.09**	**-.03**	20	.21	10	.31	**.19**	.29
11	0	0	17	n/c	n/c	n/c	5	.05	16	.57	**.20**	.44	12	.13	14	.54	**.22**	.45
12	4	.06	15	**.01**	**-.25**	**-.15**	4	.04	17	.29	**-.02**	**.15**	9	.09	16	.33	**.09**	.25
13	4	.06	15	**-.05**	**-.14**	**-.11**	12	.12	11	**-.10**	**-.22**	**-.20**	17	.18	11	**.22**	**.20**	**.24**
14	6	.09	9	**-.06**	**-.20**	**-.15**	11	.11	14	**-.11**	**-.21**	**-.19**	16	.17	12	**.15**	**.06**	**.13**
15	6	.09	10	**-.06**	**-.20**	**-.15**	12	.12	12	**-.01**	**-.20**	**-.13**	15	.16	13	**.18**	**.18**	**.21**
16	6	.09	11	**-.06**	**-.20**	**-.15**	12	.12	13	**-.02**	**-.20**	**-.13**	10	.10	15	**-.06**	**-.01**	**-.04**
17	4	.06	16	**-.05**	**-.14**	**-.11**	9	.09	15	**-.12**	**-.18**	**-.18**	5	.05	17	**-.04**	**.04**	**-.01**
Rho:			.76						.90						.93			

TABLE 3 continued

(b) Digit Span Forward

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
1	62	.87	2	**.03**	**.14**	**.03**	97	.96	2	**.08**	**.13**	**.13**	92	.96	2	**.14**	**.19**	**.19**
2	63	.89	1	.29	**.20**	.28	98	.97	1	**.05**	**.03**	**.05**	93	.97	1	**.16**	**.21**	**.21**
3	47	.66	4	.29	.32	.35	90	.89	4	**.23**	**.20**	.25	91	.95	3	**.15**	**.13**	**.16**
4	53	.75	3	**14**	**.21**	**.20**	91	.90	3	**.21**	**.21**	.24	90	.94	4	**.17**	**.15**	**.19**
5	7	.10	6	.31	**.24**	.31	35	.35	6	.29	**.15**	.25	38	.40	6	**.16**	**.16**	**.18**
6	18	.25	5	**.23**	**.08**	**.17**	64	.63	5	**.20**	**.09**	**.17**	71	.74	5	.25	**.23**	.28
7	3	.04	7	.36	**.16**	.29	20	.20	7	**.18**	**.01**	**.11**	33	.34	7	.26	**.19**	.26
8	1	.01	8	**.24**	**.25**	.28	12	.12	8	**.07**	**.11**	**.10**	18	.19	8	**.21**	**.09**	**.18**
9	0	0	9	n/c	n/c	n/c	5	.05	9	**.08**	**.12**	**.11**	6	.06	10	**.11**	**.12**	**.14**
10	0	0	10	n/c	n/c	n/c	2	.02	10	**.05**	**.12**	**.10**	8	.08	9	**.18**	**.15**	**.19**
11	0	0	11	n/c	n/c	n/c	0	0	11	n/c	n/c	n/c	2	.02	11	**-.01**	**.11**	**.06**
12	0	0	12	n/c	n/c	n/c	0	0	12	n/c	n/c	n/c	1	.01	12	**-.05**	**-.02**	**-.04**
13	0	0	13	n/c	n/c	n/c	0	0	13	n/c	n/c	n/c	0	0	14	n/c	n/c	n/c
14	0	0	14	n/c	n/c	n/c	0	0	14	n/c	n/c	n/c	1	.01	13	**.04**	**.18**	**.12**
Rho:			.93						.96						.98			

TABLE 3 continued

(c) Digit Span Backward

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*R*PS	*r*FS	*f*	*p*	R	*r*VS	*R*PS	*r*FS	*f*	*p*	R	*r*VS	*R*PS	*r*FS
1	55	.78	2	.38	.28	.37	96	.95	2	**.23**	.40	.37	95	.99	1	**.04**	**-.07**	**-.02**
2	59	.83	1	.44	.40	.48	97	.96	1	**.18**	.35	.32	95	.99	2	**.03**	.32	**.20**
3	32	.45	3	.39	.33	.41	72	.71	3	**.16**	.38	.32	82	.85	3	**.23**	**.18**	**.24**
4	26	.37	4	.37	**.24**	.34	70	.69	4	**.01**	**.16**	**.11**	72	.75	4	**-.01**	**.09**	**-.01**
5	9	.13	5	**.16**	**.27**	**.25**	29	.29	6	**.02**	**.09**	**.06**	50	.52	5	**.09**	**.02**	**.06**
6	6	.09	6	**.08**	**.02**	**.05**	31	.31	5	**.01**	**.21**	**.13**	48	.50	6	**.19**	**-.01**	**.11**
7	0	0	8	n/c	n/c	n/c	4	.04	8	**.13**	**.07**	**.12**	14	.15	8	.25	**.13**	**.22**
8	1	.01	7	**.11**	**.07**	**.10**	6	.06	7	**-.03**	**.08**	**.03**	16	.17	7	**.01**	**.13**	**.08**
9	0	0	9	n/c	n/c	n/c	0	0	10	n/c	n/c	n/c	3	.03	10	**-.01**	**.04**	**.02**
10	0	0	10	n/c	n/c	n/c	2	0	9	**.06**	**-.03**	**.01**	4	.04	9	**-.02**	**-.06**	**-.04**
11	0	0	11	n/c	n/c	n/c	0	0	11	n/c	n/c	n/c	1	.01	11	**.04**	**.18**	**.12**
Rho:			.96						.95						.99			

hard). Only two items are of moderate difficulty for G3 and fully 69% fall within the extreme deciles of difficulty. For Digits Backward (Table 3c), for G1, the majority of these items (57%), (G2: 0%; G3: 1 item) correlate significantly with the Verbal Score; although 43% (G2: 3 items; G3: 1 item) also correlate significantly with the Performance Score. Thus, Digits Backward is also a relatively weak discriminator between the verbal and performance constructs of the WISC-R. As with Digits Forward, Digits Backward shows quite a bit of consistency in item order (G1: *rho* .96; G2: *rho*=.95; G3: *rho*=.99), with no item misplaced by more than two ranks. Only two items are moderately difficult for G1, & G2 (G3: 1 item) and another two fall within the lower decile of difficulty (too hard), for G1, while no items are too easy. A large increase in difficulty is apparent after the second pair of items (4 digits) for G1. The first item-pair (#1 & 2) is too easy and the fifth item-pair (#9 & 10) is too difficult for G2; for G3, only two items (#5 & 6) are of moderate difficulty whereas 45% fall within the extreme deciles of relative difficulty.

Picture Completion Subtest. An examination of Table 4a reveals that, for G1, 58% (G2: 54%; G3: 46%) of the scorable items correlate significantly with the Performance Score, and only 19% (G2: 15%; G3: 15%) correlate significantly with the Verbal Score, suggesting moderately good discrimination for the Performance Scale for G1 and 2, whereas it is not a particularly strong measure of any of the three WISC-R constructs for the older Inuit. The Picture Completion subtest also is quite well ordered with respect to item difficulty for G1 (*rho*=.96; G2: *rho*=.88; G3: *rho*=.95) with only five of 26 items misplaced by more than two ranks from standard for G1: #12 (*"man"*), #13 (*"door"*) and #18 (*"scissors"*) are easier, and #6 (*"mirror"*) and #10 (*"dresser"*) are more difficult. For G2, however, 13 items are misplaced by more than two rank positions from standard: #11 (*"belt"*), #12, 13, 18, 24 (*"telephone"*), #25 (*"profile"*), all easier, and #5 (*"cat"*), #8 (*"elephant"*), #9 (*"ladder"*), #10, 15 (*"girl running"*), #21 (*"cow"*), and #23 (*"house"*), all more difficult. For G3, Picture Completion is really only relatively well ordered for the first few items; 10 of the 26 scorable subtest items are misplaced by more than two ranks from standard. Items #8, 10, 15, 21, and 22 (*"thermometer "*) are more difficult; items #11, 14 (*"playing card"*), #17 (*"boy"*), #24, and 25 are easier. For G1, Picture Completion includes a relatively large portion of moderately difficult items (42%; G2: 19%; G3: 1 item) and items within the extreme deciles of difficulty (38%; G2: 43%). For G3, 54% are too easy and only 1 item is too hard. Thus, this subtest includes a relatively large portion of items (15%) that are clearly too easy for even the very young Inuit children (G1). A moderate increase in difficulty is apparent beginning with item #20 for G1; item #21 (*"cow"*), for G2 and G3.

Picture Arrangement Subtest. Table 4b reveals that, for G1, fully 92% (G2: 75%; G3: 75%) of the scorable items correlate significantly with the Performance Score, although 50% (G2: 25%; G3: 0%) of scorable items also correlate significantly with the Verbal Score; suggesting that for the young Inuit (G1) this subtest contains a relatively large "verbal" component to performance. Picture Arrangement appears to be a reasonably good measure of the WISC-R performance construct for the older groups. This subtest is also reasonably well ordered with respect to relative difficulty for the children (G1: *rho*=.90; G2: *rho*=.97; G3: *rho*=.93); with only two items misplaced by more than two ranks from standard for G1 (G2=0 items; G3: item #9); #9 (*"chase"*) is easier, while #6 (*"rush"*) is more difficult. Fully 50% of Picture Arrangement items are of moderate difficulty for G1 (G2: 42%; G3: 3 items only) and two items (#1 & 2) are clearly too easy (G2: items #1, 2, 3; G3: items 1 to 4).

Block Design Subtest. Table 4c shows that, for G1, 78% (G2: 73%; G3: 64%) of

TABLE 4

Frequency and Relative Probability of Correct Response, Item Order Based on Probability of Correct Response, and Correlation of Items With Verbal Score, Performance Score, and Full Score for Three Age Pools

(a) Picture Completion Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS	f	p	R	rVS	rPS	rFS
1	69	.97	1	**.27**	**.20**	**.26**	101	1.00	1	n/c	n/c	n/c	95	.99	1	**.11**	**.12**	**.13**
2	68	.96	3	.29	**.21**	.28	101	1.00	2	n/c	n/c	n/c	95	.99	2	**.11**	**.12**	**.13**
3	69	.97	2	**.22**	**.27**	.28	101	1.00	3	n/c	n/c	n/c	95	.99	3	**.11**	**.12**	**.13**
4	68	.96	4	**.25**	.34	.34	101	1.00	4	n/c	n/c	n/c	95	.99	4	**.11**	**.12**	**.13**
5	60	.85	5	**.25**	**.12**	**.20**	91	.90	9	.26	**.23**	.28	93	.97	5	**.15**	**.03**	**.11**
6	46	.65	11	.39	.49	.51	95	.94	5	**.13**	**.17**	**.17**	90	.94	6	.31	.28	.34
7	57	.80	6	.34	.42	.44	92	.91	8	**-.06**	**.16**	**.07**	90	.94	7	**.11**	**.13**	**.14**
8	48	.68	9	**.16**	.28	**.26**	88	.87	11	**.15**	.28	.26	89	.93	9	**.13**	.28	**.24**
9	48	.68	10	.35	.60	.55	81	.80	13	**-.07**	.34	**.17**	89	.93	10	**.11**	.27	**.21**
10	39	.55	14	**.10**	.41	.31	79	.78	15	**.13**	.31	.26	82	.85	15	**.17**	**.20**	**.21**
11	45	.63	12	**.13**	**.26**	**.23**	94	.93	6	**.04**	.25	**.18**	90	.94	8	**.18**	**.19**	**.22**
12	54	.76	7	**.12**	.32	.26	93	.92	7	**.05**	.34	.24	89	.93	11	**.03**	**.20**	**.13**
13	50	.70	8	**.23**	.40	.37	90	.89	10	**.19**	**.22**	**.20**	89	.93	12	**.10**	**.12**	**.13**
14	40	.56	13	**.17**	.47	.37	82	.81	12	**-.04**	.26	**-.14**	89	.93	13	**.05**	**.23**	**.16**
15	26	.37	17	**.22**	.46	.40	59	.58	20	.25	.28	.32	76	.79	18	.26	.35	.35
16	28	.39	16	**.12**	.38	.30	64	.64	17	**.09**	**.16**	**.15**	78	.81	17	**.21**	.28	**.28**
17	24	.34	18	**.03**	.42	.28	68	.67	16	**.19**	.47	.39	87	.91	14	.32	.55	.50

TABLE 4 continued

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
18	30	.42	15	**.14**	.47	.36	80	.79	14	**.07**	.32	**.23**	82	.85	16	**.11**	.36	.27
19	19	.27	19	**.03**	**.25**	**.17**	60	.60	19	**.05**	.39	.27	67	.70	20	**.02**	.34	**.20**
20	8	.11	20	.41	.28	.39	62	.61	18	**.11**	.31	.25	70	.73	19	**.21**	.33	.31
21	5	.07	22	**-.08**	**.04**	**-.02**	11	.11	24	**.11**	**.18**	**.17**	16	.17	24	**.15**	.30	.26
22	6	.09	21	**.21**	.41	.36	13	.13	23	**.10**	**.20**	**.18**	15	.16	25	**.08**	**.21**	**.16**
23	1	.01	25	**.15**	**.08**	**.13**	8	.08	26	.28	.32	.35	18	.19	23	.41	.46	.50
24	2	.03	24	**.16**	**.06**	**.12**	23	.23	21	**.22**	**.22**	.26	27	.28	21	**.12**	**-.01**	**.07**
25	3	.04	23	**-.16**	**.11**	**-.01**	15	.15	22	**.16**	.31	.28	21	.22	22	**.19**	.39	.33
26	1	.01	26	**.08**	**-.01**	**.04**	10	.10	25	.25	.36	.36	8	.08	26	**.13**	**.15**	**.16**
Rho:			.96						.88						.97			

TABLE 4 continued

(b) Picture Arrangement Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
1	64	.90	2	**.27**	**.21**	**.27**	100	.99	1	**.09**	**.18**	**.16**	96	1.00	1	n/c	n/c	n/c
2	65	.92	1	**.16**	.40	.33	97	.96	2	.24	**.23**	.27	96	1.00	2	n/c	n/c	n/c
3	54	.76	3	.35	.50	.49	92	.91	3	**.17**	**.11**	**.16**	93	.97	3	**.08**	.25	**.19**
4	39	.55	5	**.17**	.49	.40	88	.87	4	**.22**	.31	.31	92	.96	4	**.13**	.30	**.24**
5	47	.66	4	.41	.59	.58	78	.77	5	**.17**	.31	.29	82	.85	5	**.13**	**.21**	**.19**
6	25	.35	9	**.25**	.41	.38	55	.55	8	**.14**	.46	.36	69	.72	8	**.19**	.38	.32
7	27	.38	8	.43	.58	.58	68	.67	6	.28	.55	.49	74	.77	6	**.15**	.32	.27
8	31	.44	7	.36	.42	.45	45	.45	9	**.13**	.32	.27	62	.65	9	**.21**	.36	.32
9	33	.47	6	.34	.68	.60	57	.56	7	.39	.48	.51	74	.77	7	**.20**	.41	.35
10	10	.14	11	**.27**	.34	.35	38	.38	10	**.16**	.58	.44	47	.49	10	**.15**	.48	.36
11	11	.16	10	.40	.40	.46	18	.18	11	**.18**	.33	.30	31	.32	11	**.24**	.31	.32
12	10	.14	12	**.21**	.41	.36	13	.13	12	**.14**	.25	**.23**	20	.21	12	**.22**	.36	.33
Rho:			.91						.97						.94			

TABLE 4 continued

(c) Block Design Subtest

Item	AGE 7 & 8 YEARS (N=71)						AGE 10 & 11 YEARS (N=101)						AGE 13 & 14 YEARS (N=96)					
	f	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS	*f*	*p*	R	*r*VS	*r*PS	*r*FS
1	70	.99	1	**.14**	**.26**	**.23**	101	1.00	1	n/c	n/c	n/c	96	1.00	1	n/c	n/c	n/c
2	65	.92	2	**.21**	**.26**	**.27**	99	.98	2	**.03**	**.19**	**.13**	96	1.00	1	n/c	n/c	n/c
3	62	.87	3	.30	.34	.37	99	.98	3	**.13**	**.12**	**.15**	94	.98	3	**.07**	**.24**	**.17**
4	37	.52	5	.30	.56	.50	83	.82	5	**.10**	.37	.28	92	.96	4	**.09**	**.16**	**.15**
5	38	.55	4	.28	.66	.55	85	.84	4	**.08**	.45	.32	87	.91	5	.26	.40	.38
6	25	.35	6	**.24**	.64	.52	76	.75	6	**.21**	.56	.46	87	.91	6	.26	.38	.37
7	8	.11	7	**.17**	.39	.33	40	.40	7	**.18**	.58	.45	70	.73	7	**.20**	.55	.43
8	8	.11	8	**-.03**	.38	**.21**	40	.40	8	.33	.61	.56	60	.63	8	**.10**	.54	.36
9	6	.09	9	**-14**	.34	**.14**	28	.28	9	.34	.65	.59	60	.63	9	**.13**	.60	.41
10	0	0	10	n/c	n/c	n/c	10	.10	10	**.16**	.45	.37	19	.20	11	**.14**	.40	.31
11	0	0	11	n/c	n/c	n/c	4	.04	11	.28	.41	.40	22	.23	10	.27	.52	.45
Rho:			.98						.99						.99			

the scorable Block Design items correlate significantly with the Performance Score, whereas only 33% (G2: 27%; G3: 27%) of scorable items correlate significantly with the Verbal Score. Block Design appears to be a relatively good measure of the WISC-R performance construct with the children. The obtained item order for this subtest is also highly consistent with the standardized order (G1: *rho*=.98; G2: *rho*=.99; G3: *rho*=.99); with no items misplaced by more than two ranks. While, for G1, 33% (G2: 2 items; G3: 3 items) of items fall within the moderate difficulty range, another 33% (G2: 45%) fall within the extreme deciles of difficulty. For G1, a moderate increase in difficulty is apparent with item #4 (no block outlines in model), and again with item #7. For G2, Block Design shows a relatively stable and modest gradient of increasing item difficulty, while for G3, a sharp increase is apparent at item #10 while the first six items are clearly too easy.

Across Age Trends

Overall, a number of trends are apparent in WISC-R item-total correlations, relative difficulty, and ordering across age for the Inuit sample.[5],[6] With respect to item-total correlations, the proportion of scorable Verbal Scale items that correlate significantly with the Verbal Score appears quite stable at approximately 69-73% across age. Similarly, the proportion of scorable Performance Scale items that correlate significantly with the Performance Score also stays relatively constant at approximately 65-73% across age. However, the proportion of scorable Verbal Scale and Performance Scale items that correlate strongly (>.34) with *both* the Verbal Score and Performance Score decreases steadily with increasing age. Thus, generally, WISC-R test items appear to tend toward increasing differentiation with respect to one or the other of the WISC-R scales with increasing examinee age. With respect to item-total test correlation, the majority of test items correlate significantly with Full Score, but there is a general trend toward a slight decrease in number of scorable items with significant item by Full Score correlations with increasing age—a trend shown most strongly in Arithmetic, Digit Span, and Picture Completion.

Overall, the item ordering for subtests as given in the *WISC-R Manual* (Wechsler, 1974) is relatively consistent with an order based on presently obtained relative item difficulties for all three age pools. Only 11-16% of all scorable test items are misplaced by more than two ranks from standard in any age group. The two subtests that consistently contain the greatest relative proportion of significantly misplaced items are Picture Completion and Comprehension; whereas Arithmetic, Digit Span and Block Design contain no items that are significantly misplaced. Only 17 items (#8 in Information, #3 in Similarities, #9, 10 in Vocabulary, #11, 12, 15, 16 in Comprehension, and #10, 11, 12, 13, 15, 18, 21, 24, 25 in Picture Completion) are significantly misplaced in the same direction for two or more of the three age groups. These items need to be further examined for possible cultural bias and possibly modified, replaced, or reordered within their respective subtests for future use with Inuit children.

Of possibly greater import to the use of the WISC-R with Inuit children is the finding that only 20-24% of Verbal Scale scorable items and 10-43% of Performance Scale scorable items fall within the moderate range of relative difficulty while fully 39-51% of Verbal Scale items and 32-51% of Performance Scale items fall within the extreme deciles of relative difficulty and fail to add significantly to test variance. The proportion of scorable items in the moderate range stays quite stable across age for the Verbal Scale, but decreases steadily with increasing age for the Performance Scale. For the Performance Scale, a sharp de-

crease in the proportion of extremely difficult items (15% to 2%) and a sharp increase in the proportion of too easy items (17% to 49%) is evident with increasing age. Thus it appears that the Performance Scale suffers from a relatively low performance ceiling with these children.

Conclusions

Overall, the WISC-R appears to be a reasonably consistent measure of the Wechsler verbal, performance, and global abilities constructs with Inuit children. The vast majority of test items demonstrated adequate and appropriate discrimination between the Wechsler verbal and performance constructs. As with the original WISC-R standardization sample, the Arithmetic and Digit Span subtests appear to measure some other, orthogonal, construct with the Inuit sample. This is particularly true of the older Inuit whose responses to these two subtests show particularly weak relationships to their responses on the other seven subtests. Possibly this other construct is similar to the *Freedom from Distractability* construct suggested by Kaufman (1975) based on his factor analysis of the WISC-R standardization data, although a recent study by Stewart and Moely (1983) suggests that Arithmetic and Digit Span may be measuring quite different cognitive processes and, therefore, may not be a unitary third WISC-R construct. Little evidence was found to suggest the necessity for major item ordering changes or content revision beyond those already made to "Canadianize" the WISC-R; although a small number of items were consistently misplaced across age levels.

Of somewhat greater concern is the present finding, similar to Seyfort et al., (1980), that only approximately one-third of scorable WISC-R items at any age were of moderate difficulty (*P*-value = .30 to .70) for the Inuit sample while roughly another third fell within the extreme deciles of relative difficulty and failed to add significantly to total test response variance. Although no similar item difficulty data has been published for the original WISC-R standardization sample to permit direct comparison, it would seem likely that these figures for the Inuit sample represent a significant increase in overall WISC-R difficulty compared to that shown by majority children. If this is indeed the case, Canadian indigenous children would tend to be evaluated on the basis of fewer test responses and, therefore, WISC-R measurement reliability would be compromised to some extent. With the Inuit children, an overall increase in the relative difficulty of the WISC-R (particularly the Verbal Scale items) may be a function of any of the factors commonly given as reasons for Native disadvantage on intellectual assessment instruments. That is,

> bilingual language orientations and nonverbal childrearing styles, coupled with nonoral-visual modality preferences, may reduce the Indian child's accumulation of standard English language content and efficient non-Indian linguistic processing capabilities. An emphasis upon social and emotional stimuli and relationships rather than upon transfer of abstract information and factual knowledge may be reinforced through traditional cultural values, low socioeconomic status, and lack of school success. In addition, an extremely high incidence of otitis media and consquent hearing loss in Indian groups may affect psycholinguistic development, reducing auditory-verbal reception, association, sequencing, and expressive abilities. (McShane & Plas, 1984, p. 70)

A strategy of renorming the WISC-R with Native children may operate to adjust overall scores in a manner to compensate for increased test difficulty due to socioeconomic disadvantage and/or cultural differences, but fails to resolve a possible measurement problem and may actually decrease the test's criterion predictive validity.

Clearly, a renorming strategy should go hand in hand with studies of test criterion validity and measurement reliability. Further research is necessary to evaluate the WISC-R's test-retest reliability and predictive validity with Inuit children. At present, the use of appropriately renormed standardization tests and a "social utilities" approach (Novick & Petersen, 1976; Petersen & Novick, 1976) in selection and placement decisions appears to be the most politically and socially palatable methodology for ensuring fairness to individuals and social responsiveness to minority and/or culturally-different groups (Reschley & Sabers, 1979).

Notes

1. Requests for reprints should be addressed to Robert F. Mulcahy at the Department of Educational Psychology, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5. The collection of the data upon which the present paper is based was funded by the Department of Education, Government of the Northwest Territories. The production of this paper was partially supported by the Social Sciences and Humanities Research Council of Canada through a Doctoral Fellowship to the first author. The authors wish to acknowledge and thank David W. Matheson, Department of Educational Psychology, University of Alberta, for his invaluable assistance.
2. Mulcahy, R., & Watters, B. (1983). *Report on Phase 1: Kitikmeot & Keewatin Region Norming Project*. Unpublished document. Available from the Department of Education, Government of the Northwest Territories, Yellowknife, N.W.T., Canada.
3. For Picture Completion, items remained unchanged; however, only for the first nonresponse, the examiner said: "Look carefully and tell me what is missing." For Similarities, an item was added at the end of the subtest, ("How are a kilogram and a meter alike?"), but was not included in the scoring. Also, when a child responded simply "yes" or "no," the examiner asked: "Tell me how?" Arithmetic item #10 was replaced with: "If one candy bar costs $.08, how much will three candy bars cost? " and item #11 replaced "supermarket" with "the Bay," (contraction for the Hudson's Bay Company stores common in the Canadian North). For Vocabulary, examiners were told to use up to two queries if, after one query, the child gave only part of the answer and had stopped.
4. Inspection of the tables will provide information on new placement order in terms of difficulty level for the Inuit children.
5. Without reporting mean "IQ" values for the Inuit children, which are misleading given that the test items show bias for the sample, it will be noted that, using the original Wechsler norms, none of the 366 children scored above 119 on Full Scale IQ (i.e., 0% in the superior to very superior range) and 32.24% scored below IQ=70, in contrast to Wechsler's reported percentages of 9.7% and 2.2%, respectively, for the normative group.
6. It will be recalled that all item response frequencies and item difficulty gradients involved a 0-1 scoring system. As noted earlier, three of the subtests are unaffected. For the other six, there is some reduction of response variance and lowered item-total correlations; however ordinal consistency, a major concern of the paper, is increased. These limitations must be considered in interpreting the findings.

References

Anastasi, A. (1976). *Psychological Testing*. New York, NY: MacMillan.

Dash, U., Dennis, S., Mueller, H.H., Mancini, G., Snart, F., & Short, R. (1983). WISC-R subtest variability in a sample of clinic-referred Canadian children. *Canadian Journal of Behavioral Science, 15*, 211-227.

Kaufman, A. (1975). Factor analysis of the WISC-R at 11 age levels between 6 1/2 and 16 1/2 years. *Journal of Consulting and Clinical Psychology, 43*, 135-147.

Kleinfield, J. (1973). Intellectual strengths in culturally different groups. *Review of Educational Research, 43*, 341-359.

MacArthur, R. (1968). Assessing intellectual potential of Native Canadian pupils: A summary. *The Alberta Journal of Educational Research, 14*, 115-122.

Marx, R. (1984). Canadian content and the WISC-R Information subtest. *Canadian Journal of Behavioral Science, 16*, 30-36.

McShane, D., & Plas, J. (1984). The cognitive functioning of American Indian children: Moving from the WISC to the WISC-R. *School Psychology Review, 13*, 61-73.

Mishra, S. (1982). The WISC-R and evidence of test bias for Native-American Navajos. *Psychology in the Schools, 19*, 458-464.

Novick, M., & Petersen, N. (1976). Towards equalizing educational and employment opportunity. *Journal of Educational Measurement, 13*, 77-88.

Peters, H. (1976). The validity of the WISC-R. *Canadian Journal of Behavioral Science, 8*, 414-417.

Petersen, N., & Novick, M. (1976). An evaluation of some models for culture fair selection. *Journal of Educational Measurement, 13*, 3-29.

Reschly, D., & Sabers, D. (1979). Analysis of test bias in four groups with the regression definition. *Journal of Educational Measurement, 16*, 1-16.

Sandoval, J. (1979). The WISC-R and internal evidence of test bias with minority groups. *Journal of Consulting and Clinical Psychology, 47*, 919-927.

Sandoval, J., & Miille, M. (1980). Accuracy of judgments of WISC-R item difficulty for minority groups. *Journal of Consulting and Clinical Psychology, 48*, 249-253.

Sandoval, J., Zimmerman, I., & Woo-Sam, J. (1983). Cultural differences on WISC-R verbal items. *Journal of School Psychology, 21*, 49-55.

Seyfort, B., Spreen, O., & Lahmer, V. (1980). A critical look at the WISC-R with Native Indian children. *Alberta Journal of Educational Research, 26*, 14-24.

Spreen, O., & Tryk, H. (1970). WISC Information subtest in a Canadian population. *Canadian Journal of Behavioral Science, 2*, 294-298.

St. John, J., Krichev, A., & Bauman, E. (1976). Northwestern Ontario Indian children and the WISC. *Psychology in the Schools, 13*, 407-411.

Stewart, K., & Moely, B. (1983). The WISC-R third factor: What does it mean? *Journal of Consulting and Clinical Psychology, 51*, 940-941.

Taylor, L., & Skanes, C. (1976). Cognitive abilities in Inuit and white children from similar environments. *Canadian Journal of Behavioral Science, 8*, 1-8.

Vance, H., Gaynor, P., & Coleman, M. (1977). Item analysis of the WISC-R. *Psychology in the Schools, 14*, 132-139.

Vernon, P. (1974). WISC-R. *Canadian Psychology Association Bulletin, 4*, 8-9.

Vernon, P. (1976). Modifications of the WISC-R for Canadian use. *Canadian Psychology Association Bulletin, 6*, 4-5.

Vernon, P. (1977). Final report on modifications of WISC-R for Canadian use. *Canadian Psychology Association Bulletin, 7*, 5-7.

Wechsler, D. (1974). *WISC-R manual.* New York, NY: Psychological Corporation.

Wersh, J., & Briere, J. (1981). WISC-R subtest variability in normal Canadian children and its relationship to sex, age and I.Q. *Canadian Journal of Behavioral Science, 13*, 77-86.

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 37-50*

BARRY J. FRASER
Western Australian Institute of Technology

WAYNE W. WELCH
University of Minnesota

and

HERBERT J. WALBERG
University of Illinois at Chicago

Using Secondary Analysis of National Assessment Data to Identify Predictors of Junior High School Students' Outcomes

Data from a national sample of 2,025 13-year-olds participating in the National Assessment in Science in 1981-82 was used to test a model of educational productivity involving factors including ability, motivation, quantity and quality of instruction, and the environments of the home and class. When a set of 11 individual predictors was regressed on achievement and attitude, the educational productivity model was generally supported. With predictor variables controlled for one another, student achievement was related to ability, motivation, science teaching budget, years of science, amount of homework, class environment, home environment, amount of television viewing (negative association), gender, and race. Student attitudes were found to be positively related to ability, motivation, attitude to the teacher, amount of homework and class environment, and negatively to the amount of television viewing.

The aim of the study was to use data from a large national random sample of students in the U.S.A. to investigate the influence on cognitive and affective outcomes of a set of productive factors in learning (including ability, motivation, quantity and quality of instruction, and the environments of the home and class).

Background and Theoretical Framework

According to Walberg's theory of educational productivity (Walberg, 1981, 1983, 1984a,b; Horn & Walberg, 1984: Walberg, Tsai, & Harnisch, 1984), nine factors require optimization to increase student achievement of cognitive and affective

Dr. Fraser is Head of the School of Curriculum Studies, and Director of the Science and Mathematics Education Centre at the Western Australian Institute of Technology and his interests include classroom environment, educational evaluation, and science education.

Dr. Welch is a Professor of Educational Psychology at the University of Minnesota and the Director of the Minnesota Evaluation Center and his interests include educational evaluation, national assessment, and science education.

Dr. Walberg is a University Research Professor in the College of Education at the University of Illinois at Chicago and his interests include educational psychology, educational evaluation, and productivity research.

outcomes. These factors are the student aptitudinal variables of: (1) *ability* or prior achievement, (2) *age*, (3) *motivation* or self-concept as indicated by personality tests or willingness to persevere on learning tasks; the instructional variables of: (4) *quantity* of instruction, (5) the *quality* of the instructional experiences; and educationally stimulating psychological aspects of the: (6) *home* environment, (7) the *classroom* or school environment, (8) the *peer group* environment and (9) the *mass media* (especially television).

Compared with these nine factors, economic, political, and sociological characteristics are less relevant to learning because their influences are less alterable, direct, and observable; they are not substitutes for the nine factors, but more distant forces which can support or interfere with them. The nine productive factors largely were identified from syntheses of about 3,000 individual studies of factors related to student learning (see Walberg, 1984a,b). Some of these syntheses have involved relating student outcomes to motivation (Uguroglu & Walberg, 1979), quantity of instruction (Frederick & Walberg, 1980), home environment (Iverson & Walberg, 1982), classroom environment (Haertel, Walberg, & Haertel, 1981), and leisure-time television viewing (Williams, Haertel, Haertel, & Walberg, 1982).

Few prior intensive experiments and quasi-experiments are national in scope, and most analyze only one or two of the factors and sample limited populations within a school or community. They are often strong on observational technique, measurement, verification, and random assignment to treatments (in short, internal validity), but they are often weak in generalizability or external validity since they do not sample rigorously from large, well-defined populations. Survey research has complementary strengths and weaknesses: it often draws large, stratified, random samples of national populations and measures more factors, but sacrifices internal validity since the factors are usually measured cross-sectionally and perhaps superficially with only a few items. Also survey research can control statistically to some extent for multiple causes and can be more causally convincing than quasi-experiments controlled only for one or two covariates.

The complementarity of intensive and extensive studies, however, is important. In principle, consistent, powerful effects should emerge consistently from either form of research as well as from case studies. Robust findings give a more solid basis for educational policy and practice. In fact, approximately 70 syntheses of several thousand intensive studies (Walberg, 1984b) support the consistently positive influences or correlations of the nine productivity factors or their more specific aspects. The expected negative association of achievement with hours of television viewing beyond ten per week is the exception (see Williams, et al., 1982).

In addition, a synthesis of nine regression studies of extensive survey data, involving 15,802 13- and 17-year-old students tested in mathematics, science, social studies, and reading by the National Assessment of Educational Achievement, supports the sign of these findings (Walberg, 1984b). Although the correlations of the factors with achievement and subject-matter interest as learning outcomes varied from -0.45 to +0.68, 83 (or 91 percent) of the 91 correlations were in the expected direction. Moreover, when the factors were controlled for one another in multiple regressions, 58 (or 91 percent) of the 64 coefficients carried the expected sign. Although the magnitudes of the correlations depended slightly on the type of learning outcome (achievement or attitude), subject matter, and productivity factor, they averaged +0.19, and most ranged between +0.05 and +0.40. The standardized regression weights, moreover, were nearly all in the expected direction and not significantly different across age, subject matter, type of outcome, and produc-

tivity factor. The analyses, however, revealed collinearity among the factors; students advantaged on one also tended to be advantaged on the others.

The results suggest further that no single factor is dominant in determining learning outcomes and that a one standard-deviation (SD) increment in any factor (with other factors held constant) is associated with a 0.05 SD increment in learning from, for example, the 50th to the 52nd percentile, which represents a trivial advantage. Raising the complete set of factors one SD, however, is associated with a 17-percentile increment from the 50th to the 67th percentile. To take a more extreme example, an increment of two SDs from the 50th to the 98th percentile on all the factors is associated with an increment of 32 percentile points from the 50th to the 82nd percentile of learning.

Prior syntheses of intensive and extensive studies thus suggest that no single factor can serve as a panacea. Improving all the factors, using scarce resources of human time and effort as efficiently as possible, would seem a more advisable policy than improving only one, which may lead to diminishing returns. Most of the evidence, however, has been assembled from studies of a decade or more ago, and it would be useful to test the generality of the results for more current data. Thus the purpose of the present study was to compare regressions of National Assessment in Science data recently collected in 1981-82 with previous regressions and syntheses of smaller-scale experimental and quasi-experimental studies.

National Assessment in Science

The National Assessment of Educational Progress (NAEP) was established in the U.S.A. in 1969 to assess periodically students' knowledge of various school subject areas. National assessments in science were conducted by NAEP in 1969-70, 1972-73, and 1976-77. But, because of legislative decisions and financial constraints, the National Institute of Education postponed the next fully-fledged NAEP science assessment until the late 1980s. Many science educators were concerned that this hiatus would permit emerging problems to go unchecked, and so the National Science Foundation funded a national assessment which would fill this void. Consequently, during 1981 and 1982, a project under the direction of Wayne Welch at the University of Minnesota involved the collection of data on many variables including student cognitive and affective outcomes and characteristics of the home, the community, and the school (Hueftle, Rakow, & Welch, 1983; Welch, 1985). This national assessment of educational progress in science involved a national random sample of approximately 18,000 students of ages 9, 13, and 17 in about 700 schools in the U.S.A. Examples of some of the uses of this data base include examining secondary school science enrollments in the United States (Welch, Harris, & Anderson, 1984), inequities in opportunities for computer usage in schools (Anderson, Welch, & Harris, 1983), and perceptions of secondary school students toward women in science (Welch, Rakow, & Harris, 1984).

Because the main purposes of this and previous NAEP surveys are to assess the status of scientific literacy among various groups, and to compare this with images assessed in 1976-77, many analyses simply focus on the proportion of respondents in different subgroups answering each item correctly. For example, Rakow, Welch, and Hueftle (1984) report this type of data and provide the interesting conclusion that, whereas 9-year-olds increased a small amount in their science achievement over the past five years, 13-year olds showed essentially no change while 17-year-olds experienced a decline. But, because of the large number of individual test items answered by large numbers of students, the National Assessment in Science data provide an extremely useful basis for performing secondary analyses which interrelate variables (often defined in terms of subsets of individual items). In fact,

provision of a data base for science education research was a recognized aim of the 1981-82 assessment, and so, the survey included many school and student background characteristics. For example, some secondary analyses of the National Assessment in Science data already completed have explored questions such as the stability of attitudes to science between junior and senior high school (Hofstein & Welch, 1984) and the influence of class attitudes and teacher image on student outcomes (Tamir, Welch, & Rakow, in press).

Operationalizing Measures of Outcome Criteria and Productivity Factors

As the National Assessment in Science data included items which could be interpreted as measures of most of the factors in Walberg's educational productivity model, secondary analyses were conducted for the purpose of probing the validity of the model. The first step involved using the data base to operationalize two measures of student outcomes and at least one measure of seven of the nine productive factors in the model. Appendix A provides detailed descriptions and operational definitions of each of the variables involved, including scale reliability estimates where appropriate.

A cognitive achievement measure consisted of 49 multiple-choice items covering content topics, inquiry skills, and understanding of societal issues (alpha reliability=0.80). An attitudinal outcome measure consisted of 19 Likert-type items assessing opinions about the value of science and willingness to solve social problems (alpha reliability=0.68). Ability was assessed with a self-report item asking about students' previous grades in school. Age was excluded because the sample consisted only of 13-year-olds and therefore exhibited very little variability. Motivation was assessed by eight items asking about the frequency of voluntary participation in science-related activities (alpha reliability=0.78). Quality of instruction was assessed by two different variables, namely, the average science teaching budget per pupil (as reported by the school principal) and students' attitudes toward their teacher's enthusiasm, and so forth (assessed by five Likert-type items with an alpha reliability of 0.66). The quantity of instruction also was assessed by two variables, namely, the number of years of science taken over the previous two years and the average number of hours of homework per day. The class environment was measured by six items asking students how they felt during science classes (alpha reliability=0.68). The home environment was assessed in terms of the higher of the ratings for father's and mother's education coded on a scale of one to six. The peer group environment was not assessed. The mass media environment was assessed in terms of the number of hours of television watched during the previous day. In addition to the above factors in the productivity model, the present study involved the two extra variables of gender and race (white vs. non-white) because they had been found to be good predictors of learning among science students in the present sample. Of course, should these extra variables of gender and race prove to be significant independent predictors of learning, this would suggest an omission in the original productivity model.

These descriptions of variables show that several of the productive factors are incompletely, crudely, or unreliably indexed by the questionnaire items available as part of the National Assessment in Science. As this leads to underestimates in the magnitude of associations between learning and productive factors, it is likely that future studies involving better measures of variables would yield even stronger relationships than those emerging in this study.

Sample

The National Assessment in Science selected respondents of ages 9, 13, and 17 using a stratified, two-stage probability sample design to ensure proportional representation by regions, males and females, different racial groups, and size and type of community. The first stage of the sampling involved dividing the U.S.A. into primary sampling units consisting of geographic groupings with a minimum population size requirement. These primary sampling units were stratified according to four regions (Northeastern, Southeastern, Central, and West) and five sizes of community strata (ranging from large cities with more than 200,000 population to extreme rural areas of less than 10,000). The sample was drawn with probabilities proportional to size to represent all regions and community sizes. Oversampling of low-income and rural areas ensured adequate representation of these groups. In the second stage of sampling, 125 schools were randomly chosen with probabilities proportional to the size of the school. Finally, a random sample of approximately 16 students was selected from each school.

Although the total sample consisted of approximately 18,000 students in about 700 schools throughout the U.S.A., the group of almost 8,000 13-year-olds tested in 298 schools was selected for consideration for the present secondary analyses. Because the study involved a matrix sampling plan whereby different students completed different parts of the total test battery, the secondary analyses were restricted to a subgroup of 2,025 13-year-olds who had completed all items in one test booklet which made up the variables in the productivity model. In an attempt to ensure uniformity of testing conditions, test administrators visited each school to give the instruments using a timed and paced audiotape.

Analyses and Results

Table 1 provides some descriptive statistics for each of the outcomes measures and productivity factors. This information includes the mean and standard deviation of each variable and the intercorrelations among variables for the present sample of 2,025 students. For the two categorical variables of gender and race (which involved binary coding with 1=male, 0=female, 1=white, and 0=non-white), the mean simply represents the proportion of males or whites present in the total sample. For example, Table 1 shows that the sample was made up of 48% of males and 79% of whites.

Table 1 shows that some of the stronger simple correlations with cognitive achievement were for ability (0.30), class environment (0.25), and race (0.28). The larger correlations between attitudes and productive factors emerged for instruction (0.43), attitude to the teacher (0.35), and class environment (0.36).

Table 2 shows the results obtained when the whole set of 11 predictors was regressed on achievement and attitude. Use was made of a "backwards elimination" method in which variables were progressively dropped from the full regression model containing all predictors until all variables remaining had regression weights significantly different from zero ($p<0.05$). The main statistics reported for each productive factor in Table 2 are the raw regression weight, b, and the result of a t-test of whether the magnitude of the regression weight is greater than zero. Table 2 also shows that the multiple correlation for the whole of a set of predictor variables was 0.52 for achievement (27% of variance accounted for) and 0.51 for attitude (26% of variance accounted for). The total amounts of variance are not large and suggest the need in future national assessments of better and more sensitive measures of many of the educational productivity factors.

TABLE 1

Number of Items, Alpha Reliability (Where Appropriate), Mean, Standard Deviation, and Intercorrelations for all Achievement Measures and Productive Factors for 13-Year-Olds (N = 2,025)

Variable	No. of Items	Alpha Reliability	Mean	SD	Intercorrelations*												
					Achievement	Attitude	Ability	Motivation	Budget	Teacher	Years of Sc.	Homework	Class Env.	Home Env.	TV	Gender	Race
Outcomes																	
Achievement	49	0.80	28.56	6.84	--	29	30	23	09	12	23	16	25	18	-.09	12	28
Attitude	19	0.68	66.27	8.19		-	18	43	-.01	35	01	19	36	09	-10	03	-04
Ability	1		5.80	1.60			-	15	04	11	15	18	25	16	-.04	-10	09
Motivation	8	0.78	2.70	0.88				-	-05	38	02	18	46	08	-07	15	-11
Quality of instruction																	
Science teaching budget/ pupil	1		4.91	8.86					-	-08	11	01	-02	02	-04	00	11
Attitude to teacher	5	0.66	3.64	0.68						-	01	13	39	04	-05	02	-05
Quantity of instruction																	
Years of science	1		1.63	0.64							-	12	06	07	02	-04	11
Homework	1		1.53	0.76								-	13	11	-09	-08	06
Class environment	6	0.68	3.40	0.67									-	14	-05	10	-06
Home environment	1		4.88	1.11										-	-07	02	06
TV	1		2.80	2.09											-	09	-08
Gender	1		0.48													-	02
Race (white, non-white)	1		0.79														

* Decimal points have been omitted.

TABLE 2

Raw Regression Weights and t Ratios For Each Significant Productive Factor For 13-Year-Olds (N=2,025)

Productive Factor	Achievement		Attitude	
	b	t	b	t
Ability	0.81	9.28**	0.38	3.70**
Motivation	1.09	6.36**	2.51	12.00**
Quality of instruction				
Science teaching budget/ pupil	0.04	2.45*		
Attitude to teacher			2.16	8.37**
Quantity of instruction				
Years of science	1.66	8.00**		
Hours of homework per day	2.05	2.27*	4.30	4.04**
Class environment	1.14	5.17**	1.70	6.07**
Home environment	0.51	4.23**		
Hours of TV per day	-0.17	-2.70**	-0.21	-2.76**
Gender	1.56	5.80**		
Race (white, non-white)	4.12	12.44**		
Multiple R	0.52		0.51	

*p<0.05, **p<0.01

The number of significant independent predictors for the achievement outcome was 10 for the cognitive outcome and six for the affective outcome (see Table 2). When controlled for all other variables, ability, motivation, science teaching budget, the number of years of science taken over a two-year period, the amount of homework, the classroom environment, the home environment, the amount of television viewing (negative direction), gender, and race were all significantly related to science achievement. For the attitude outcome, the variables linked with science attitudes, when controlled for other variables, were ability, motivation, attitude to the teacher, amount of homework, classroom environment, and amount of television viewing (negative direction).

Raw regression weights are reported in Table 2 because they indicate the change in the number of points on an outcome measure associated with a one-unit increment in each independent variable when all other independent variables are held constant. Of course, corresponding standardized beta weights could be calculated readily using the standard deviations listed in Table 1. For example, Table 2 shows that, for cognitive achievement and with other factors held constant, an increase of one hour per day in the time spent on homework is associated with 2.05 points increase, while a decrease of one hour of television viewing per day is associated with an increase of 0.17 of a point. Also for the binary variables coded 0 and 1, regression weights can be interpreted directly as group differences. For instance, on cognitive achievement, boys scored 1.56 points higher than girls and whites scored 4.12 points higher than non-whites when the other productive factors were held

constant. That is, with other factors fixed, white boys scored 5.68 (about five-sixths of a standard deviation) higher than non-white girls on the achievement test.

It is interesting to note that several relatively unalterable factors—specifically, ability, home environment, gender, and race—are significant predictors of student cognitive achievement. It is also somewhat disappointing to find that the two measures of quality of instruction—namely, the science teaching budget per pupil (marginally significant) and students' attitudes to their teacher (nonsignificant)—are relatively weak predictors of achievement. Nevertheless, six relatively alterable factors were still found to be significant independent predictors of science students' cognitive achievement; these are motivation, science teaching budget, the number of years of science, the amount of time devoted to homework, the class environment, and the amount of television viewing (negative relationship).

Five of the factors found to be significant independent predictors of cognitive achievement also were found to be significant independent predictors of student attitudes; these are ability, motivation, amount of homework, class environment, and amount of television viewing. But, additionally, attitude to the teacher proved to be a significant independent predictor of students' attitudes. On the other hand, five factors important in predicting cognitive achievement—namely, science teaching budget, years of science, home environment, gender, and race—all turned out to be unimportant in predicting affective outcomes when other factors are controlled. Still, there are a number of alterable factors which teachers can concentrate on if they wish to attempt to improve students' attitudes to science.

The positive influence of the home environment on cognitive achievement is consistent with syntheses of prior research studies (Graue, Weinstein, & Walberg, 1983; Iverson & Walberg, 1982). Similarly, the negative association between achievement and amount of television viewing is consistent with Williams, Haertel, Haertel, and Walberg's (1982) research synthesis which revealed a deleterious effect beyond 10 hours of viewing per week. The finding of links between achievement and the two measures of quantity of instruction in the present study (namely, number of years of science and amount of homework) is consistent with Frederick's (1980) review of research on instructional time. The magnitudes of the correlations emerging in the present study, however, are considerably smaller than the correlation of 0.73 found between mathematics achievement and the number of semesters of mathematics instruction in Welch, Anderson, and Harris' (1982) analysis of the 1977-78 NAEP data of 17-year-olds. These differences in the sizes of the correlation between achievement and quantity of instruction for science and mathematics might be explained partly by the fact that there are more opportunities for out-of-school learning in science (e.g., museums, zoos, television) than in mathematics.

It is noteworthy that the nature of the classroom psychosocial environment emerged as a significant predictor of both science achievement and attitudes to science when other factors were held constant. The simple outcome-class environment correlations of 0.25 for achievement and 0.36 for attitudes to science are among the larger correlations found for any of the productivity factors in Table 1. These findings replicate individual studies of the effects of classroom environment on student achievement (Fraser & Fisher, 1982) and attitudes (Fraser, 1981) and are consistent with Haertel, Walberg, and Haertel's (1981) meta-analysis.

Conclusion and Discussion

The results show that a number of factors, previously revealed to be consistent correlates of achievement in syntheses of small-scale bivariate research, were found

also to be significant predictors of student outcomes when mutually controlled for each other in a large national survey. This suggests that national achievement and attitudes are jointly influenced by a number of factors rather than by a single dominant one. Moreover, the results of these secondary analyses generally support the validity of the model of educational productivity (although two factors not included in the model, namely, gender and race, proved to be significant independent predictors of cognitive achievement).

The present research has identified four productivity factors which predict cognitive achievement and which are relatively unalterable by the school; these are ability, home environment, gender, and race. Of these four variables, ability is the only factor also associated with student attitude. Nevertheless, there are five relatively school-alterable factors (namely, motivation, quality of instruction, quantity of instruction, class environment, and amount of television viewing) which were significant independent predictors of both achievement and attitude. It is encouraging to find that there are numerous school-alterable factors which teachers can work on if they wish to improve the cognitive and affective performance of their junior high school students.

This study highlights the problem arising from the mismatch between the purposes of the primary and secondary analyst. The national assessment data contained only crude and insensitive measures of several of the productivity factors. This is not surprising, of course, since the original survey was not designed for the secondary analyst's purpose of testing the productivity model. Because of the generality of the factors in the productivity model and the model's potential for improving learning, it would be highly desirable for future surveys of the National Assessment of Educational Progress to incorporate high-quality measures of the productivity factors.

Despite some acknowledged shortcomings, this paper still attests to the potential usefulness of secondary analyses of data from large-scale assessments (Fraser & Tobin, 1985). Clearly, there were important advantages in using the National Assessment in Science data base in the present study when compared with collecting primary data to investigate the same questions. These include the reduction in response burden on students and schools, the low cost of data, the quality of the large, national, random sample, the generalizability of results, and the possibility of future replications. Although a potential problem commonly facing secondary analysts is difficulty in understanding and using the data because of its complexity or inadequate documentation, fortunately, this was overcome with the National Assessment in Science data because of thorough documentation, careful sampling and test administration, and the use of a variety of quality checks at every stage of assessment and analysis.

Given the limitations in the quality of some of the measures of the productivity factors, more sophisticated analyses than those reported here would not be justified. Nevertheless, in future tests of the productivity model using more sensitive measures, benefit is likely to be obtained by using more sophisticated path analysis models, such as LISREL (Joreskog & Sorbom, 1984), in which some of the productive factors can be treated as endogenous interviewing variables. As well, analyses could try to establish directional links between achievement, attitude, and productivity factors. This use is illustrated by Riley and Schibeci (1983), who used LISREL with the 1976-77 NAEP science data to show that attitude influenced achievement, while background factors similar to the productivity factors affected both achievement and attitude.

Future secondary analyses of data bases containing more sensitive measures also might usefully examine non-linear relationships between productivity factors and outcomes. Walberg (1981) has noted that there are likely to be "diminishing returns" associated with raising several of these factors (e.g., quantity of instruction). Amount of homework provides a good example of a variable which is likely to bear a non-linear relationship with achievement. Similarly, Williams, Haertel, Haertel, and Walberg's (1982) meta-analysis has shown that the negative relationship between television viewing and achievement was only evident in the range between two and eight hours (so that a deleterious effect was not present under two hours per day and over eight hours per day).

There remain numerous important research questions which cannot be explored using currently available data bases of the National Assessment of Educational Progress. Consequently, a desirable thrust for national assessments in the future would be the inclusion of measures and variables which would allow researchers to investigate the reasons why productivity factors promote student achievement and attitude. For example, what is the mechanism by which home environment, gender, and race influence student achievement? Clearly these variables are proxies for factors in the home, the school or elsewhere which are the real causes of differential performance. If the data from national surveys could be utilized to pursue these important questions, it is likely that researchers would uncover information to guide our attempts to reduce the large differences in achievement evident among students differing on such dimensions as home background, gender, and race.

The findings emerging from the present analyses refute several claims made in relation to school learning and assumptions made in educational research. First, although much prior research has been bivariate in nature, this study clearly illustrates that no single factor alone can produce marked increases in learning. Improving all the productive factors using scarce resources, including time and effort, as efficiently as possible would seem a more advisable policy than improving only one. Second, the results are at variance with another commonly held view that school learning is too complex to be understood in terms of a relatively small number of underlying factors. The third notion dispelled by the present findings is that the only important factors in predicting student outcomes are those that cannot be altered by teachers or the school.

References

Anderson, R.A., Welch, W.W., & Harris, L.J. (1983). Computer inequities in opportunities for computer literacy. *The Computing Teacher, 11* (8), 10-12.

Fraser, B.J. (1981). Predictive validity of an individualized classroom environment questionnaire. *The Alberta Journal of Educational Research, 27*, 240-251.

Fraser, B.J., & Fisher, D.L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal, 19*, 488-518.

Fraser, B.J., & Tobin, K. (Eds.). (1985). *Secondary analysis and large-scale assessments* (Faculty of Education Research Seminar and Workshop Series). Perth: Western Australian Institute of Technology.

Frederick, W.C. (1980). Instructional time. *Evaluation in Education, 4*, 117-118.

Frederick, W.C., & Walberg, H.J. (1980). Learning as a function of time. *Journal of Educational Research, 73*, 183-194.

Graue, M.E., Weinstein, T., & Walberg, H.J. (1983). School-based home instruction and learning: A quantitative synthesis. *Journal of Educational Research, 76*, 351-360.

Haertel, G.D., Walberg, H.J., & Haertel, E.H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal, 7*, 27-36.

Hofstein, A., & Welch, W.W. (1984). The stability of attitudes towards science between junior and senior high school. *Research in Science and Technological Education, 2*, 131-138.

Horn, E., & Walberg, H.J. (1984). Achievement and interest as functions of quantity and level of instruction. *Journal of Educational Research, 77*, 227-232.

Hueftle, S., Rakow, S., & Welch, W. (1983). *Images of science: A summary of results from the 1981-82 National Assessment in Science.* Minneapolis: Minnesota Research and Evaluation Center, University of Minnesota.

Iverson, B.K., & Walberg, H.J. (1982). Home environment and school learning: A quantitative synthesis. *Journal of Experimental Education, 19*, 144-151.

Joreskog, K.G., & Sorbom, D. (1984). LISREL VI: *Analysis of linear structural relationships by the method of maximum likelihood.* Chicago: National Educational Resources.

Rakow, S., Welch, W., & Hueftle, S. (1984). Student achievement in science: A comparison of national assessment results. *Science Education, 68*, 571-578.

Riley, J.P., & Schibeci, R., (1983, April). *Influence of student background on science attitudes and achievement.* Paper presented at Annual Meeting of American Educational Research Association, Montreal.

Tamir, P., Welch, W.W., & Rakow, S.J. (in press). The influence of science class attitudes and teacher image on student outcomes. *Journal of Research and Development in Education.*

Uguroglu, M.E., & Walberg, H.J. (1979). Motivation and achievement: A quantitative synthesis. *American Educational Research Journal, 37*, 375-390.

Walberg, H.J. (1981). A psychological theory of educational productivity. In F. Farley, & N. Gordon (Eds.). *Psychology and education.* Berkeley, CA: McCutchan.

Walberg, H.J. (1983). Scientific literacy and economic productivity in international perspective. *Daedalus, 112*, 1-28.

Walberg, H.J. (1984a). Improving the productivity of America's schools. *Educational Leadership, 41*, 19-30.

Walberg, H.J. (1984b). Synthesis of research on teaching. In M.C. Wittrock (Ed.), *Third handbook of research on teaching.* Washington, DC: American Educational Research Association.

Walberg, H.J., Tsai, S.L., & Harnisch, D. (1984). High school productivity in ten countries. *Journal of Educational Research.*

Welch, W.W. (1985). Secondary analysis and national assessment in science in the U.S.A.. In B.J. Fraser and K. Tobin (Eds.). *Secondary analysis and large-scale assessments* (Faculty of Education Research Seminar and Workshop Series). Perth: Western Australian Institute of Technology.

Welch, W.W., Anderson, R.E., & Harris, L.J. (1982). The effects of schooling on mathematics achievement. *American Educational Research Journal, 19*, 145-153.

Welch, W.W., Harris, L.J., & Anderson, R.E. (1984). How many are enrolled in sciences? *The Science Teacher, 51*, 14-19.

Welch, W.W., Rakow, S.J., & Harris, L.J. (1984, April). Women in science: Perceptions of secondary school students. Paper presented at Annual Meeting of National Association for Research in Science Teaching. New Orleans, LA.

Williams, P.A., Haertel, E.H., Haertel, G.D., & Walberg, H.J. (1982). The impact of leisure-time television on school learning: A research synthesis. *American Educational Research Journal, 19*, 19-50.

Appendix A

Descriptions and Operational Definitions of Variables

Achievement Outcome

Forty-nine multiple-choice items covering science content topics (e.g., energy, life, changing aspects of the earth), inquiry skills (including experimentation, data interpretation, measurement, problem definition and solution), understanding of the role of theories in science and an appreciation and understanding of societal issues (including health, resource management, nutrition, safety, and the costs and benefits associated with applied science and technology). Maximum score=49. Alpha reliability=0.80.

Attitude Outcome

Nineteen Likert-type items assessing students' opinions about the usefulness and value of science and willingness to solve social problems (e.g., by turning off lights when they are no longer needed). Alpha reliability=0.68. Each item is responded to on a five-point scale. Items are coded as:
5 = Always or Strongly agree or Definitely yes
4 = Often or Agree or Probably yes
3 = Sometimes or No opinion or No response
2 = Seldom or Disagree or Probably not
1 = Never or Strongly disagree or Definitely not
Some items are scored in the opposite direction. Maximum score=95.

Ability

A self-report item asking students to describe grades in school so far coded as:
8 = Mostly A
7 = About half A and half B
6 = Mostly B
5 = About half B and half C
4 = Mostly C
3 = About half C and half D
2 = Mostly D
1 = Mostly below D

Motivation

The mean of eight items asking how often science-related activities (e.g., reading science articles in magazines, watching science shows on TV, going to hear people give talks on science) have been done when not required for science classes. Alpha reliability=0.78. Coded as:
5 = Often
4 = Sometimes
3 = No opinion or No response
2 = Seldom
1 = Never

Science Teaching Budget

Science teaching budget (in dollars) per student obtained from a principal's questionnaire reporting the school's total instructional budget, the percentage devoted to science and the number of students in the school.

Attitude to Teacher

The means of five Likert-type items requesting students' opinion about their present or most recent science teacher. The items ask whether the teacher "really likes science," "wants students to point out mistakes. . .," "makes science exciting," "is enthusiastic," and "is willing to share opinions. . .." Alpha reliability=0.66. Coded as:
5 = Strongly agree
4 = Agree
3 = No opinion
2 = Disagree
1 = Strongly disagree

Quantity of Science Instruction

Based on two items asking students if they are taking a science course currently and whether they were taking one at the same time during the previous year. Coded as:
0 = taking a science course in neither year
1 = taking a science course in one of the years
2 = taking a science course in both years

Homework

An item asking students to indicate the average amount of time per day spent on homework (all school subjects). Coded as:
0 = None
1 = Less than one hour per day
2 = Between one and two hours per day
3 = More than two hours per day

Class Environment

The mean of six items asking students if science classes made them feel "uncomfortable" (reverse scoring), "curious," "stupid" (reverse scoring), "confident," "successful" and "unhappy" (reverse scoring). Alpha reliability=0.68. Except for items with reverse scoring, coding is:
5 = Always
4 = Often
3 = Sometimes
2 = Seldom
1 = Never

Home Environment

The higher of the ratings for father's and mother's education coded as:
1 = Did not complete eighth grade
2 = Completed eighth grade, but did not go to high school

3 = Went to high school, but did not graduate from high school
4 = Graduated from high school
5 = Some education after graduation from high school
6 = Graduated from college

TV

An item asking students how many hours of TV they watched during the previous day. Coded as:
1 = Watched one hour or less
2 = two hours
3 = three hours
4 = four hours
5 = five hours
6 = six hours or more

Gender

Coded as:
1 = Male
0 = Female

Race

Racial background with blacks, native Americans, Hispanics, and Asians included in the non-white category. Coded as:
1 = White
0 = Non-white

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 51-65*

ALLAN J. MACDONELL
University of New Brunswick

and

WILFRED B.W. MARTIN
Memorial University of Newfoundland

Student Orientations to School Rules[1]

This paper uses data from a survey questionnaire administered to approximately 8,000 high school students in 55 randomly selected schools in Newfoundland and Labrador to identify the social background, organizational, and student characteristics which are associated with student orientations to school rules. Analyses of students' perceptions of number of rules, necessity of rules, necessity of following rules, impossiblity of obeying rules, frequencies of breaking school rules, whether it is right or wrong to break school rules, and frequencies of individual and group approaches in teacher-student interactions concerning the implementation of rules, indicate that while certain variables in each of the three categories (social background, organizational and student characteristics) are more important than others, student characteristics are more often related to students' perceptions of different dimensions of school rules than are social background and organizational characteristics.

Rules, as a means of legitimizing rational-legal authority, are the core of the bureaucratic phenomenon in contemporary society. In line with Weber's (1964, pp. 333-334) ideal type bureaucratic model, a system of rules covering the rights and obligations of its members is generally seen as an important characteristic of the school organizational entity. By way of setting the stage for the present focus on factors relating to student orientations to rules, it is appropriate to give a brief overview of particular organizational dimensions of the school.

Organizational Dimensions of the School

Two interrelated topics of long-standing interest in the organization of the school have particular relevance to the present research. One is the issue of student alienation and the other is student involvement in decision-making in the school.

Dr. Macdonell is a Professor in the Department of Sociology, University of New Brunswick with interests in the sociology of education with particular reference to higher education.
Dr. Martin is a Professor in the Department of Educational Foundations, Memorial University of Newfoundland, with interests in the sociology of education with particular reference to micro level processes in the school.

Using the concepts of "status maintenance" and "behavior control" to describe students' perceptions of six dimensions of bureaucratic structure, Anderson (1970) found that students' alienation is negatively related to their perceived status maintenance. The implication is that student alienation may increase as bureaucratic efficiency increases.

Several Canadian studies have focused on student involvement in the decision-making processes in the school (Treslan, 1969; Wood, 1977; Alexander & Farrell, 1975; Hoffman, 1971; and Probert, 1971). These studies show a lack of student involvement in the decision-making processes in the school and a concomitant student desire for greater participation in this dimension of school life. The question of student orientation to school rules has been alluded to in these studies of students' perceptions of their involvement in decision-making in the school. For example, in their study of students in Ontario, Alexander and Farrell (1975) found that students think they have a minimal influence in most areas of decision-making in the school. An Alberta study by Treslan (1969) found that, with few exceptions, teachers and administrators are perceived to be making all decisions concerning rules and regulations in the school.

Even though it has been argued that the nature of teaching is such that "rule domination is much less compatible with the performance of school activities than it is with . . . the management of industrial production of tangible goods and services" (Dreeben, 1970, p. 47), rules are one of the dominating features of many high schools. There is usually a host of rules which attempt to govern student behavior from the moment they arrive on the school grounds, to their entering the school and their moving around in the corridors, to their visits to the washroom and other specifically defined areas in the school, including directions for a multitude of classroom activities and potentially problematic situations. In fact, schools often develop extensive lists of rules, and put considerable effort into communicating them through student handbooks, posting lists on bulletin boards and on classroom walls, and by repeating them during the teaching-learning processes.

Given the widespread use of bureaucratic rules in the school in general, and in teacher-student interaction in particular, it is in many ways surprising that research on different aspects of this phenomenon is so sparse. A host of rule-related issues has yet to be studied and more fully understood in the quest to get at the complexities of the interrelationships among such things as student attitudes and behaviors, organizational features of the school and societal structures and processes. In attempting to understand these complexities, it is seen that one of the most salient features of the formal organization of the school is the rules, particularly those prescribing the rights and obligations of the students. The present research focuses on student orientations to school rules. More specifically, the intent is to see what social background characteristics of students, organizational characteristics of schools and student characteristics are related to and associated with student orientations to school rules.

Data Collection and Analysis

Data on which this publication is based were collected as part of a larger study of the schooling experience of high school students in Newfoundland and Labrador. The larger study used three data collection techniques: semi-case study, taped interviews, and a survey questionnaire. The questionnaire was administered to 7,848 high school students from 55 schools representing 24 school boards in Newfoundland and Labrador. Actually, 58 schools were selected by means of a random sample and stratified on the basis of community size. But the questionnaire was not administered in two of these schools because the school boards having jurisdic-

tion over them did not grant the researcher permission to approach them concerning this study. A third school in the sample had to be omitted because its principal refused to let the questionnaire be administered in his school. The sample of almost 8,000 high school students is evenly divided along gender lines and proportional representation from each high school grade was attained.

Data collected on parts of three questions on the survey questionnaire form the basis of the present analysis of student orientation to school rules. One of these questions asked students to indicate how often (never, seldom, sometimes, often, very often) they did certain things in the school building:

- Break certain rules which are given concerning classroom behavior;
- Break certain rules which are given concerning behavior outside of the classroom, but within the school or on the school grounds.

Another question attempted to find out how often students, individually and/or in groups, approached their teachers as they attempted to get them to change their minds concerning a variety of issues in the school, including:

- The following rules (for example, dress code, hair styles, smoking, where to walk in the school corridors, what to do on the school grounds, behavior while in the classroom and behavior while in other areas of the school).

A third question asked students to indicate whether they agreed or disagreed with each of a list of 50 statements. The responses received on six of these statements form part of the data base for the present analysis. These six statements are:

- It is necessary to follow all the school rules all of the time.
- There is nothing wrong with breaking certain school rules occasionally.
- It is impossible to obey all the rules in this school.
- There are too many rules in this school.
- Most school rules are necessary.
- It is okay to break the rules at school so long as you do not get caught.

These questions provide data on students' perceptions concerning eight dimensions of rules. One is their perceptions of the number of rules in the school. Another is their perceptions of the necessity of rules. Thirdly, there is the extent of the student-perceived necessity to follow all school rules. Fourthly, the extent to which students perceive it is impossible to obey all the school rules is of interest here. Other areas of concern are students' perceptions of the frequencies with which they break classroom rules and the frequencies with which they break other school rules. A seventh dimension of rules investigated here revolves around whether it is right or wrong to break school rules, and under what circumstances it is perceived to be right or wrong to follow particular courses of action with respect to obeying rules. Finally, attention is given to the frequencies with which students perceive themselves to be using specific interaction approaches in attempting to get their teachers to change their minds concerning the enforcement of rules. Taken together these perceptions are referred to in the present paper as student orientations to school rules.

Chi square test of significance is used to isolate the relationships between specific social background, organizational, and student characteristics, on the one hand, and student orientations to rules, on the other. The social background, organizational, and student characteristics are further divided into categories of independent variables, as follows:

I. Social Background

1. Family size
   - number of children in family
2. Parental employment and occupation
   - whether or not father is employed
   - whether or not mother is employed
   - father's occupation
   - mother's occupation
3. Parental education
   - secondary education (grade completed)
   - post-secondary education (attended university and/or trades college)
4. Size of home community

II. Organizational Characteristics

1. Type and size of school
   - grades in school (high school grades, junior high and high school grades, all grade school)
   - number of classes in high school grades
   - number of students in high school
2. School organizations
   - student representative council
   - a system of prefects
   - school clubs (organizations)
   - average size of school clubs
3. School board
   - denominational schools (Integrated, Roman Catholic and Pentecostal)

III. Student Characteristics

1. Individual characteristics
   - gender
   - age
   - ethnicity
2. Organizational membership
   - membership in student representative council
   - being a prefect
   - number of school clubs one has membership in
3. Evaluation indicators
   - grade level
   - whether or not one failed a grade
   - number of grades failed
   - number of grades repeated
   - whether or not one failed a subject
   - number of subjects failed repeated
4. Busing and mobility factors
   - whether or not one is bused to school
   - whether or not one has lived in the same community all one's life
   - numbers of years one has lived in the present community of residence

The associations between these orientations to rules, taken as dependent variables, and the specific social background, organizational and individual student characteristics as outlined above, taken as independent variables, are measured by using gamma. This test measures the relative reduction in errors of predicting order on the dependent variable. That is to say, the numerical value of gamma, disregarding whether it is positive or negative, gives the percentage of guessing error eliminated by using knowledge of an independent variable to predict order on a dependent one (Mueller, Schuessler, and Costner, 1977, pp. 187-188; 217-219).

Findings

In line with the categories of independent variables included in the present analysis, the findings are presented by giving separate attention to social background variables, organizational characteristics, and student characteristics.

Social Background

Summaries of chi squares and gammas for the relationships and associations between different dimensions of student orientations to school rules and each of the social background variables considered in the present study are given in Table 1. In connection with employment and occupation, it is interesting to note that none of the dimensions of student orientations is significantly related to whether or not father is employed, and only one of three dimensions is significantly related to whether or not mother is employed. The associations between these dimensions of student orientations to school rules and father/mother employment are also low. The highest gamma is only .09. Six of the dimensions of student orientations to school rules are significantly related to father's occupation and four are significantly related to mother's occupation when occupations are categorized along the Blishen and McRoberts (1976) scale of social classes. Again, the associations are generally low, with only two of a total of twenty gammas above .10 (Table 1).

None of the dimensions of student orientations to rules is significantly related to the father's schooling or to the mother's schooling, or to the question of whether or not the mother went to trades school, or to whether or not the mother went to university. Only one of these dimensions is significantly related to whether or not the father went to trades school, and only two of the ten orientations are related to whether or not father went to university. The associations between student orientations to school rules and education are low, with only three of the 60 associations given in the education grouping of Table 1 being higher than gamma of .10.

Of the four categories of social background variables (family size, employment/occupation, education, and community size), community size seems to be most important in terms of number of orientations significantly related to population of community, and in terms of the strength of associations between these two variables. Seven of the ten dimensions of student orientations are significantly related to population of community, and four of the ten associations between these two variables range from gamma of .10 to gamma of .14. Even though these gamma are somewhat low, they are high in comparison to the associations between student orientations to school rules and the other social background variables considered in this analysis (Table 1).

TABLE 1

Orientations to Rules and Social Backgrounds: Summaries of Chi Squares and Gammas

Student Orientations to School Rules		Family Size			Employment/ Occupation		Education						Community Size
		Number of Children (3)	Father Employed (5)	Mother Employed (5)	Father's Occupation (4)	Mother's Occupation (4)	Father's Schooling (3)	Mother's Schooling (3)	Father's Trade School (5)	Mother's Trade School (5)	Father's University (5)	Mother's University (5)	Population of Community (3)
Perceived number of rules (1)	X^2	*			*	*							*
	Y	-.09	-.07	-.05	-.12	-.07	.07	-.00	-.01	-.02	-.05	-.05	.13
Necessity of rules (1)	X^2												
	Y	-.06	-.01	-.04	.04	-.02	.01	-.01	-.04	.02	-.10	-.07	.01
Necessity of following rules (1)	X^2	*		*	*	*					*		*
	Y	-.15	-.09	-.09	-.11	-.07	.03	.06	-.05	-.03	-.14	-.11	.11
Impossibility of obeying rules (1)	X^2												
	Y	.02	-.04	.00	.00	.03	-.01	-.03	-.01	-.03	.02	-.08	.02
Breaking classroom rules (2)	X^2				*								*
	Y	-.01	-.02	-.03	-.06	-.03	.04	.05	-.04	.01	-.05	-.06	.10

TABLE 1 continued

Student Orientations to School Rules		Family Size			Employment/ Occupation		Education						Community size
		Number of Children (3)	Father Employed (5)	Mother Employed (5)	Father's Occupation (4)	Mother's Occupation (4)	Father's Schooling (3)	Mother's Schooling (3)	Father's Trade School (5)	Mother's Trade School (5)	Father's University (5)	Mother's University (5)	Population of Community (3)
Breaking other school rules (2)	X^2								*				*
	Y	-.00	-.01	-.04	-.03	-.02	.00	.02	-.04	-.01	-.04	-.01	.07
Breaking rules occasionally (1)	X^2	*			*						*		*
	Y	.07	.06	.05	.08	.07	-.04	-.06	.02	.01	.13	.07	-.10
Breaking rules and getting caught (1)	X^2												
	Y	-.03	-.01	.03	.00	.03	.02	.04	-.01	-.05	-.01	-.07	-.05
Individual approach to interaction (2)	X^2	*			*	*							*
	Y	.10	.03	.03	.03	.04	.01	.01	-.01	.-01	.04	-.01	-.02
Group approach to interaction (2)	X^2	*			*	*							*
	Y	.10	.04	.02	.02	.03	.00	.00	-.02	-.01	.03	-.03	-.04

Note: Direction (1) = agree/disagree, (2) = never/very often, (3) = low/high, (4) = high/low, (5) = yes/no
* = $X^2 < .001$

TABLE 2

Orientations to Rules and Organizational Characteristics: Summaries of Chi Squares and Gammas

Student Orientations to School Rules		Type and Size of School			School Organizations				School Board
		Type of School (4)	Number of Classes at Grade Level (3)	Number of Students in High School Grades (3)	Student Representative Council (5)	A System of Prefects (5)	School Clubs (5)	Average Size of School Clubs (3)	School Board (Integrated Roman Catholic/ Pentecostal).
Perceived number of rules (1)	X^2	*	*	*			*	*	*
	Y	-.17	.14	.14	-.04	.01	-.15	-.18	-.20
Necessity of rules (1)	X^2								
	Y	.02	.04	.01	-.04	.08	.08	.17	-.04
Necessity of following rules (1)	X^2	*	*	*		*	*		*
	Y	-.08	-.13	.13	-.08	.13	-.20	-.01	-.15
Impossibility of obeying rules (1)	X^2		*	*				*	
	Y	-.02	-.01	.00	-.01	-.04	.03	-.12	.07
Breaking classroom rules (2)	X^2		*	*	*	*		*	
	Y	-.03	.10	.08	-.09	.09	-.03	.12	.03
Breaking other school rules (2)	X^2		*	*		*		*	*
	Y	.02	.05	.05	-.08	.10	-.06	.15	.06

TABLE 2 continued

Student Orientations to School Rules		Type and Size of School			School Organizations				School Board
		Type of School (4)	Number of Classes at Grade Level (3)	Number of Students in High School Grades (3)	Student Representative Council (5)	A System of Prefects (5)	School Clubs (5)	Average Size of School Clubs (3)	School Board (Integrated Roman Catholic/ Pentecostal).
Breaking rules occasionally (1)	X^2	*	*	*	*	*			*
	Y	.08	-.11	-.13	.12	-.13	.14	-.08	.13
Breaking rules and getting caught (1)	X^2		*	*				*	*
	Y	.04	-.09	-.08	.10	-.09	-.02	-.25	.16
Individual approach to interaction (2)	X^2	*	*	*			*		*
	Y	.13	-.07	-.08	.00	.04	.08	.03	.29
Group approach to interaction (2)	X^2	*	*	*					*
	Y	.13	-.08	-.09	-.01	.06	.05	.01	.35

Note: Direction (1) = agree/disagree, (2) = never/very often, (3) = low/high, (4) = high/low, (5) = yes/no
* = $X^2 < .001$

Organizational Characteristics

Table 2 contains summaries of the chi squares and gammas for the relationships and associations between student orientations to school rules and the variables in each of the three groupings of organizational characteristics; type and size of school, school organizations, and school board.

Student responses to five of the statements indicating their orientations to school rules are significantly related to type of school. The associations between the responses to some of these statements and type of school are noteworthy with gammas of .13 and .17. Both measures of school size (number of classes at grade level and number of students in high school grades) are important variables in that, except for one dimension, student orientations are significantly related to each of these variables. Even though a couple of the associations between these orientations and school size variables are almost nonexistent, four of the associations have gammas at .13 and .14, and other associations range from .07 to .09 (Table 2).

A survey of the four school organizations reveals that only two of the student orientation variables are significantly related to the question of whether or not there is a student representative council in the school, four of the dimensions of student orientations are related to the existence of a system of prefects, three of the student orientations are related to school clubs, and five student orientations are related to average size of school clubs. A couple of the associations between student orientations and the school organizations are noteworthy: that between necessity of following rules and school clubs—gamma .20, and between perceived number of rules and average size of school clubs—gamma .18 (Table 2).

Seven of the dimensions of student orientations to rules are significantly related to school board, and the associations between three of these orientations and school board are relatively high with gamma of .35 between group approach and school board, gamma of .29 between individual approach and school board, and gamma of .20 between perceived number of rules and school board (Table 2). That is to say, of the organizational variables considered in this analysis, school board is the most important one in terms of predicting student responses to statements indicating their orientations to school rules. According to chi square test, both school board and school size are the most important variables in the organizational category. The finding concerning the connection between school size and student orientations to school rules can be seen to be important in the continuing debate over the influence of school size on different dimensions of school life (see, e.g., Riffel, 1971; Sackney, 1976; Hickcox & Burston, 1973; and Holdaway & Blowers, 1971). Also, the present finding concerning school boards and student orientations to school rules is suggestive of the need to look at the school board as a unit of analysis in attempting to understand the schooling experiences of high school students. Others, for example, Bidwell (1974/1975) and Cooley (1983), have pointed to the need to focus on school district in analysing different features of schooling.

Student Characteristics

By summarizing the chi squares and gammas between student orientations to school rules and four groupings of student characteristics (individual characteristics, organizational membership, evaluation indicators, and busing and mobility factors), the importance of certain independent variables are highlighted while others are seen not to be significantly related to or associated with these orientations. Nine of the ten dimensions of student orientations focused on in this research are significantly related to gender. In addition, the association between

each of these nine, taken separately, and gender are, with only one exception, relatively strong indeed. For example, there is a gamma of .47 between breaking other school rules and gender, and there are other gammas of more than .40 between breaking classroom rules and gender, and between breaking rules and getting caught and gender. Gammas ranging from .30 to .40 are found in three other dimensions of student orientations to school rules and gender. The strong associations between gender and different dimensions of student orientations to school rules may be an indication that the sex-role stereotyping which was demonstrated to be a part of schooling experiences a decade ago (Lorimer & Long, 1979/1980; Pyke & Ricks, 1975; and Pyke, 1975) is still a part of, at least, the hidden curriculum of the school.

While student responses to six of the statements addressing their orientations to school rules are significantly related to age, the associations between these responses and age are low with gammas ranging from .05 to .09. Only one of the dimensions of student orientations to school rules (group approach to interaction) is significantly related to ethnicity, and the gamma of association between these two variables is only low at .08. However, the associations between two other dimensions of student orientations (necessity of rules and necessity of following rules) and ethnicity are considerably higher with gammas of .18 and .14, even though these orientations are not significantly related to ethnicity (Table 3). Given the importance of ethnicity in a wide range of schooling issues such as the dropout problem and academic achievement in Canadian society (Henry, 1973; Nagler, 1975), the present finding suggesting a low relationship between ethnicity and student orientation to school rules is interesting indeed. It challenges us to the all too often taken for granted influence of particular social demographic features on schooling.

Turning to the organizational membership grouping of student characteristics, it is noted that only one of the dimensions of student orientations to school rules is significantly related to membership in student representative council, and the association between these two variables is moderately strong—gamma .16. Whether or not one is a prefect is an important variable in studying student orientations to school rules in that their responses to eight of the ten aspects of rules are significantly related to it. Many of the associations between these orientations and this independent variable are also relatively strong. Two of the gammas are over .30 (necessity of rules, breaking rules and getting caught), and three are between .20 and .30 (necessity of following rules, breaking classroom rules and breaking other school rules). Another three gammas are at .19. It has also been found that nine dimensions of student orientations to school rules are significantly related to the number of school clubs which students are members of, and the associations between each of these dimensions as dependent variables and the number of school clubs as independent variables are generally moderately strong. With a couple of exceptions, gammas of .02 and .06, the gammas range from .11 to .23 (Table 3).

Also, in attempting to understand student orientations to school rules, all of the evaluation indicators (grade level, whether or not one failed a grade, number of grades failed, number of grades repeated, whether or not one failed a subject, number of subjects failed, and number of failed subjects repeated) have been found to be important in that student orientations to school rules are, with few exceptions, significantly related to each of these indicators. While there are some very weak, almost non-existing, associations between certain dimensions of student orientations and some of these evaluation indicators for example, gammas of .01, .02 and .03 between certain orientations and grade level, overall the gammas between these orientations and each of the evaluation indicators are moderate to high, many of

TABLE 3

Orientations to Rules and Student Characteristics: Summaries of Chi Squares and Gammas

Student Orientations to School Rules		Individual Characteristics			Organizational Membership			Evaluation Indicators							Busing and Mobility		
		Gender (Male/Female)	Age (3)	Ethnicity (English/French/Native)	Member in Student Representative Council (5)	Being a Prefect (5)	Number of School Clubs (3)	Grade (3)	Grade Failed (5)	Number of Grades Failed (3)	Number of Grades Repeated (3)	Subject Failed (5)	Number of Subjects (3)	Number of Failed Subjects Repeated (3)	Busing (5)	Same Community All Life (5)	Number of Years In Community (3)
Perceived number of rules (1)	X^2	*				*	*	*	*	*	*	*	*	*	*		
	Y	.29	-.04	-.04	-.00	-.19	.11	.14	.28	-.25	.27	.25	-.21	.19	.12	.06	-.04
Necessity of rules (1)	X^2	*				*	*		*	*	*	*	*	*			
	Y	-.39	.02	.15	.10	.31	-.14	-.04	-.13	.13	-.14	-.16	.14	-.10	.10	-.00	.04
Necessity of following rules (1)	X^2	*	*			*			*	*	*				*		
	Y	-.30	.08	.18	-.10	.22	-.06	-.17	.11	-.11	.10	.07	-.06	.04	.09	.04	.01
Impossibility of obeying rules (1)	X^2	*				*	*		*	*	*	*	*	*			
	Y	.20	-.03	.00	-.02	-.19	.14	.02	.15	-.13	.14	.22	-.18	.17	.05	.01	-.03
Breaking classroom rules (2)	X^2	*	*			*	*	*	*	*	*	*	*	*	*		
	Y	-.44	.05	.05	.03	.25	-.12	.01	-.09	.09	-.09	-.12	.10	-.11	.08	.04	-.01
Breaking other school rules (2)	X^2	*	*			*	*	*	*	*	*	*	*	*			
	Y	-.47	.09	.04	.02	.21	-.15	.02	-.15	.13	-.14	-.17	-.13	-.15	-.05	.00	.02

TABLE 3 continued

Student Orientations to School Rules		Individual Characteristics			Organizational Membership			Evaluation Indicators							Busing and Mobility		
		Gender (Male/Female)	Age (3)	Ethnicity (English/ French/Native)	Member in Student Representative Council (5)	Being a Prefect (5)	Number of School Clubs (3)	Grade (3)	Grade Failed (5)	Number of Grades Failed (3)	Number of Grades Repeated (3)	Subject Failed (5)	Number of Subjects (3)	Number of Failed Subjects Repeated (3)	Busing (5)	Same Community All Life (5)	Number of Years In Community (3)
Breaking rules	X^2	*	*			*	*	*				*					
Occasionally (1)	Y	.37	-.07	-.04	-.06	-.19	.12	.09	.03	-.01	.03	.10	-.07	.09	-.05	-.01	-.02
Breaking rules and	X^2	*				*	*	*	*	*	*	*	*	*			
getting caught (1)	Y	.43	-.05	.04	-.20	-.33	.23	.11	.33	-.30	.31	.35	-.28	.26	-.00	.03	-.04
Individual approach	X^2	*	*		*		*	*	*	*	*	*	*	*	*		
to interaction (2)	Y	-.11	.05	.09	-.16	.05	-.02	-.03	-.17	.15	-.17	-.12	.10	-.08	.03	.08	.05
Group approach	X^2		*	*				*	*	*	*	*	*	*			
to interaction (2)	Y	.02	.05	.08	-.09	.08	-.02	.01	-.11	.09	-.10	-.08	.07	-.06	.80	-.05	.05

Note: Direction (1) = agree/disagree, (2) = never/very often, (3) = low/high, (4) = high/low, (5) = yes/no
* = $X^2 < .001$

which are in the .20 to the .30 range (Table 3).

Only four of the dimensions of student orientations to school rules are significantly related to busing, and the associations between these four and busing are relatively weak with gammas of .03 to .12. None of the student orientations to schools rules is significantly related to whether or not students lived in the same community all of their lives or to the number of years one has lived in one's present community of residence (Table 3).

Implications

This study points to the need to be cognizant of theoretical issues and practical concerns in the society-school arenas as one strives to gain greater understanding of the schooling process in general and of the rule processes in particular. On a theoretical level, there is a need to distinguish between objective and subjective dimensions of the school organization. In the present study, attention is given to the subjective dimension, that is, the student orientations to school rules. The findings suggest that student orientations are significantly related to (X^2 test), and associated with (gamma), student characteristics of gender, membership in student representative councils, being a prefect and all the evaluation indicators included in this analysis. Hence, there is a need for theoretical development to find ways of explaining the saliency of these variables over the social background variables such as parental occupation and parental education which have been used traditionally to explain different aspects of the schooling experiences of high school students. In fact, the overall relationship between social background variables, on the one hand, and student orientations to school rules, on the other, is worthy of further study. It may be that the issue of the interrelationships among the processes of class, culture, and education are more complex than is often realized. The connections between these processes are, perhaps, only glossed over when the major focus is on parental education and occupation.

Turning to the practical implications of the present focus on student orientations to school rules, the suggestion here is that identifying the variables associated with these orientations is important on two accounts. One is the fact that having a greater awareness of the differences in student orientations along certain variables such as those included in the social background, organizational and student categories leads to a greater understanding of the human dimensions of rules. In other words, the factors which play important roles in students' interpretations of school rules and of the meanings they give to these rules are highlighted. These interpretations and meanings are core elements of the social structure in the day-to-day operation of the school. Another practical implication of the present approach to rules, and the findings presented herein, is the highlighting of the importance of certain parts of the school organization. For example, the finding that student orientations are significantly related to whether or not one is a prefect and to the number of clubs one has membership in may have implications for the development of these aspects of the school organization. Similarly, the significance of evaluation indicators in student orientations to rules reaffirms the importance of these parts of school life on student orientations to school rules, which in turn are intertwined with disciplinary and other teaching-learning processes in the school.

Note

1. Appreciation is hereby expressed to the Social Sciences and Humanities Research Council of Canada. Without the financial support of the Council the research on which this paper is based would not have been possible.

References

Alexander, W.E., & Farrell, J.P. (1975). *Student Participation in Decision-making.* Toronto: The Ontario Institute for Studies in Education.

Anderson, B.D. (1970). Bureaucratization and alienation: An empirical study in secondary schools. Unpublished doctoral dissertation, University of Toronto.

Bidwell, C.E. (1974/1975). The relationship between school district organization and student achievement. *Administrator's Notebook, 13*, 1-4.

Blishen, B.R., & McRoberts, H.H. (1976). A revised socioeconomic index for occupations in Canada. *Canadian Review of Sociology and Anthropology, 13*(1), 71-79.

Cooley, W.W. (1983). Improving the performance of an educational system. *Educational Researcher*, 12(6), 4-12.

Dreeben, R. (1970). *The Nature of Teaching: Schools and the Work of Teachers.* Glenview, IL: Scott, Foresman & Company.

Henry, F. (1973). *Forgotten Canadians: The Blacks of Nova Scotia.* Don Mills: Longman Canada.

Hickcox, E.S., & Burston, G. (1973). The question of size. *Education Canada, 13*(3), 41-43.

Hoffman, G. (1971). A survey of senior high school student satisfaction-dissatisfaction in an Alberta urban community. Unpublished Master's of Education thesis, University of Calgary.

Holdaway, E.A., & Blowers, T.A. (1971). Administrative ratios and organizational size: A longitudinal examination. *American Sociological Review, 36*(2), 278-286.

Lorimer, R., & Long, M. (1979/80). Sex-role stereotyping in elementary readers. *Interchange, 10*(2), 25-45.

Mueller, J.H., Schuessler, K.F., & Costner, H.L. (1977). *Statistical Reasoning in Sociology* (3rd ed.). Boston: Houghton Mifflin.

Nagler, M. (1975). *Natives without a home.* Toronto: Longman Canada.

Probert, E.G. (1971). Shared decision-making in the high school. Unpublished Master's of Education thesis, University of Alberta.

Pyke, S.W. (1975). Children's literature: Conceptions of sex roles. In R.M. Pike & E. Zureik (Eds.), *Socialization and values in canadian society.* Vol. II, pp. 5-73. Toronto: McClelland and Stewart Limited.

Pyke, S.W., & Ricks, F. (1975). Teachers push sex differences. *Special Education in Canada, 50*(1), 4-5.

Riffel, J.A. (1971). *Educational planning reexamined.* Edmonton: Human Resources Research Council of Alberta.

Sackney, L.E. (1976). The relationship between organizational structure and behaviour in secondary schools. Unpublished doctoral dissertation, University of Alberta.

Treslan, D.L. (1969). The control assembly: An administrative alternative in Canadian senior high schools. *Commonwealth Council for Educational Administration Studies in Educational Administration*, No. 16.

Weber, M. (1964). *The theory of social and economic organization.* Translated by A.M. Henderson, & T. Parsons. New York: The Free Press.

Wood, J.D. (1977). Student influence in decision-making in secondary schools. Unpublished doctoral dissertation, University of Toronto.

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 66-76*

BARRY C. CARBOL
British Columbia Ministry of Education

and

THOMAS O. MAGUIRE
University of Alberta

Benefits and Disappointments from Participation in the Alberta High School Equivalency Program[1]

The Alberta High School Equivalency Program provides adults living in Alberta with an opportunity of acquiring a high school diploma, either through successfully passing the Tests of General Educational Development (GED) or through the assessment of a variety of educationally relevant personal factors. Samples of successful GED applicants, unsuccessful GED applicants, and people who received the diploma through the alternate route were surveyed to determine their perceptions of the benefits resulting from the program and their demographic characteristics. The results suggest that personal benefits such as self-satisfaction outweighed educational and job-related benefits. There was some evidence to suggest that those who had acquired an equivalency diploma may have enhanced their chances for employment. Of the demographic variables considered, only age was found to be important in explaining the survey results.

It is estimated that between forty and fifty percent of all Canadians who are 16 years of age or older (excluding those still in school) do not possess a high school diploma (Statistics Canada, 1980). In a society in which credentials play a critical role in selection of jobs and postsecondary education, many of these people feel they are at a disadvantage when they seek employment or further education. To assist such people, most provinces have set up programs which provide opportunities for obtaining a high school diploma without the requirement of additional high school attendance. In Alberta, the program is called the Alberta High School Equivalency Program (AHSEP).

Dr. Carbol is the Director of the Examinations Branch in the British Columbia Ministry of Education with interests in the areas of student assessments and certification.

Dr. Maguire is a Professor of Educational Psychology at the University of Alberta with interests in achievement measurement and applied statistics.

The AHSEP was introduced in June 198l as a joint venture by the two departments responsible for education within the Province of Alberta, Alberta Education and Alberta Advanced Education. Alberta Education is primarily responsible for the delivery and monitoring of programs in the areas of early childhood, elementary, and secondary education while Alberta Advanced Education's primary responsibilities are in postsecondary education. At the time of its introduction the AHSEP was seen by its sponsors as a way of providing individuals, who for one reason or another have not completed their formal high school education, with an alternative way of acquiring a high school diploma.

The program provides applicants with two ways of gaining a diploma: the assessment of a variety of educationally relevant personal factors, or the writing of the Tests of General Educational Development (GED). The personal factors are the amount of formal high school instruction received by the applicant, the applicant's age, adult education courses taken, and the extent of travel, reading, and private study done. The GED Test Program, which was developed by the American Council on Education, consists of a battery of five multiple-choice tests: Writing Skills, Social Studies, Natural Sciences, Reading Skills, and Mathematics. The tests are administered by provincial departments of education in eight Canadian provinces and the Canadian territories. State educational authorities administer the tests in all 50 American states. In Alberta, the GED tests are administered four times each year at 27 locations throughout the province. A passing standard, based on a norming study carried out with Canadian high school seniors in 1980, has been established. Alberta Education granted over 5000 equivalency diplomas through the Alberta High School Equivalency Program between December 1981 and December 1983. Of these, approximately 85% were through the GED route.

The present study was undertaken to assess the validity of the claim that the AHSEP helps program participants in meeting personal, further-educational, and job-related goals. The primary purpose of the study was to survey samples of participants to determine their perceptions of the benefits resulting from participation in the program. A secondary purpose was to determine similarities and differences among selected sub-groups of those sampled in an effort to explain the sample survey results. Information related to the Alberta equivalency population was thought to be useful to Canadian adult educators and provincial education authorities responsible for equivalency programs. To date, public policy concerning high school equivalency programs has been influenced minimally by Canadian studies related to equivalency populations.

Related Literature

The empirical literature related to the outcomes of high school equivalency programs is sparse and most of it is American. Variables of particular interest have been employer and educational institution acceptance of equivalency program graduates, employment and educational benefits derived from program participation, and a variety of personal and demographic characteristics of program participants such as sex, age, previous highest level of education, reasons for taking the GED tests, and others.

Equivalency program acceptance, especially as it is related to successful completion of the GED tests, is well documented. The United States armed forces generally accept high school equivalency diplomas as meeting one of their entrance criteria (Laurence, 1983; Waller, 1976). Similarly, the American Council on Education (1982) reported that about 88% of American colleges and universities admit applicants with GED credentials.

Beusse (1974) noted that the receipt of a high school equivalency diploma through the GED Test Program offered by U.S.A. armed forces led to higher pay and more prestigious jobs in civilian life. Similar employment benefits were reported by Ayers (1980) for graduates of an American community college, high school equivalency program, by Fischer and Evanson (1979) for Adult Basic Education graduates, by Malizio and Whitney (1985) in a survey of 1500 American companies, and by Moore (1982) for successful GED graduates in Kentucky. Anecdotal accounts in the literature also speak of enhanced employment status, opportunities, and earnings (Harris, Hedman, & Horning, 1983; Witter, 1978).

Acceptance of high school equivalency credentials is not as straightforward as it might seem, however. Cervero and Peterson (1982), based on a survey of 1906 GED candidates who wrote in 1980, found discrepancies between expected training and employment benefits of the candidates and those they actually claim to have received. Of those respondents who expected that GED credentials would help them to quality for a job, 25% reported that the credential was of no assistance. This finding was consistent with the earlier research of Dhanidina and Griffith (1975) who concluded that, while the GED may be seen as an investment by participants, little evidence exists to suggest that the credential pays off in the job market. A similar discrepancy between the expectation of being admitted to an educational program on the basis of GED certification and actual acceptance on this basis was also noted by Cervero and Peterson (1982).

Benefits, other than economic, have been detected by some researchers. Improved self-image and self-satisfaction have been most prominently mentioned (Cervero & Peterson, 1982: Fischer & Evanson, 1979: Harris, Hedman, & Horning, 1983; Waller, 1976).

Studies of American equivalency programs have also considered the personal and demographic attributes of program participants. These studies have revealed that most candidates are female (Malizio & Whitney, 1981), experience a higher level of unemployment than the national average (Behal, 1983), and have between nine and ten years of formal schooling (Malizio & Whitney, 1981; Moore, 1982; Swarm, 1981). Reasons for participating in an equivalency program are generally related to a desire to improve the participant's job situation or to the educational background of the participant (Malizio & Whitney; Swarm, 1981). Characteristics such as age, sex, and highest level of education attained also influence the motivations of those seeking an equivalency diploma (Cross, 1981; Harvey, 1978; Russ-Eft & Steel, 1980). One study has suggested that age and highest level of education are uncorrelated with GED test performance (Modjeski & Michael, 1978).

The studies of high school equivalency programs in Canada have largely been carried out by sponsoring government agencies and have shown mixed results. While other provinces have equivalency programs, only in Saskatchewan and Manitoba has there been any attempt to collect evaluation evidence from program participants. Unlike Alberta, neither of these two provinces has based an equivalency diploma on the assessment of alternative, educationally-relevant experiences.

Wedel (1974), in a survey of Saskatchewan equivalency program participants, found large Canadian national firms and federal government agencies reluctant to consider equivalency credentials as equivalent to regular high school diplomas. However, more recent surveys in Saskatchewan have reported improved acceptance of the equivalency credential (Province of Saskatchewan, 1977; 1981). A survey of Manitoba employers (Manitoba Department of Education, 1981) has also suggested a reluctance on the part of some employers to accept the equivalency credential. A 1982 Manitoba Department of Education study showed that, while 58% of the

respondents expected to gain admission to post-secondary education, only 39% did. A similar finding is reported by the Province of Saskatchewan in a 1982 survey of equivalency program participants. Program participants in both Manitoba and Saskatchewan were found to be less than 25 years of age, saw the GED tests as a way of increasing job and promotion opportunities, and reported improved self-image as the greatest program benefit (Manitoba Department of Education, 1982; Province of Saskatchewan, 1982).

Method

Between December 1981 and December 1983, 4333 applicants successfully passed the GED tests in Alberta and received a high school equivalency diploma; 802 applicants did not pass the tests. During the same period, 822 applicants received the diploma through the alternate route. From each of these three populations (GED-pass, GED-fail, and non-GED) random samples of 200 were selected from lists and files provided by Alberta Education.

Following a field trial of questionnaire items involving eight program applicants, two survey instruments were developed (one for the GED groups and a slightly different one for the non-GED group). These instruments addressed the following questions:

1. What are the educational, vocational, and demographic characteristics of program applicants?
2. Why did people apply to the program?
3. What educational and employment benefits were anticipated?
4. What benefits were realized?
5. How do the perceptions relate to group membership?

The instruments were mailed with a personalized covering letter and a stamped, addressed, return envelope to the individuals selected for study. Each instrument was numbered to assist in the mailing of a reminder letter to non-respondents. The purpose of the numbering was explained in the covering letter. A reminder letter was mailed one month later to the non-respondents. Of the 600 questionnaires mailed, 75 or 12.5% could not be delivered. The return rates, which are shown in Table 1, compare favorably with that of Cervero and Peterson (1982) who reported a rate for their survey of 24% for their survey.

Data analysis of the surveys was undertaken in three stages. First, response frequencies for each of the items in the two instruments were determined. Second, items to which large numbers of individuals across the three groups responded were subjected to cross-tabulation analyses to investigate relationships between and among the demographic and group membership variables and the perceptions of respondents concerning program expectations and benefits. Chi square tests of independence were used to determine the statistical significance of the relationships. The final stage of the data analysis involved a study of the relationship of the demographic characteristics of the GED groups and performance on the GED tests. One-way multivariate analysis of variance was used to test the significance of the relationships.

Results

Background Characteristics

Sex, age, and highest grade summaries are reported in Table 2. It is interesting to note that the GED-fail group was significantly younger ($p<.05$) than either of

TABLE 1

	Population Size	Sample	Deliverable	Returned	Percentage Returned (Ret.÷Del.)
GED-pass	4333	200	168	113	67%
GED-fail	802	200	167	77	46%
non-GED	822	200	190	133	70%

TABLE 2

Characteristics of Equivalency Program Applicants

Characteristic	Group			
	GED-pass	GED-fail	non-GED	Total
Sex				
Percentage Male	46.9	46.8	48.9	47.7
Percentage Female	53.1	53.2	51.1	52.3
Mean Age in Years	32.7	28.3	31.6	31.2
Mean Highest Grade Completed	10.0	9.8	11.3	10.5

the other two groups. This finding suggested that individuals who were closer in time to having attended school were also the least successful in terms of the GED experience.

A significant difference among the three groups was also found for the highest-grade completed variable ($F[2,320]$ 69.71, $p<.001$). Post-hoc Scheffe contrasts of group means revealed significant differences between the non-GED group and the two GED groups ($p<.05$).

In general, the AHSEP participants were found to be similar to individuals who have participated in other equivalency programs. In particular, those who had written the GED tests displayed characteristics similar to GED candidates from other states and provinces. For example, the mean grade level completed by the Alberta sample was found to be grade 10, a level widely reported in the literature for different GED populations (Moore, 1982; Malizio & Whitney, 1981; Swarm, 1981).

Reasons for Applying to the Program

Most applicants, regardless of the group considered, applied to the program to improve their educational opportunities or to improve their chances of getting a better job. Individuals in the non-GED group cited "personal reasons" more often

than those in the other groups as a rationale for applying to the equivalency program. These findings are similar to those reported elsewhere (Cervero & Peterson, 1982; Province of Saskatchewan, 1982; Malizio & Whitney, 1981; Swarm, 1981).

It was not possible to distinguish the responses of subjects to questions dealing with their reasons for applying to the program based upon the highest grade they had completed. The age category to which program applicants belonged, however, seemed to influence the reasons they gave for applying to the program. In general, younger applicants were more interested in job and educational opportunities while older applicants were participating to a greater extent for personal reasons. Respondents who were less than 40 years of age were more likely than those over 40 to have applied to the program to improve their educational opportunities ($X^2[6,n=323]=13.54$, $p<.05$) while those in the 19 to 25 age group were more likely than any of the other groups to have participated to improve their chances of getting a better job ($X^2[6, n=323]=17.59$, $p<.01$). Divergent findings concerning the reasons that people of various age groups give for wanting to have an equivalency diploma were not unexpected; Harvey (1978), Russ-Eft and Steel (1980), and Cross (1981) have all found differential motivation in adult education programs to be a function of age.

Expectations and Benefits

Discrepancies between expectations and benefits were found for all three groups. This finding is again consistent with similar findings reported in the literature (Cervero & Peterson, 1982; Malizio & Whitney, 1981; Province of Saskatchewan, 1982; Swarm, 1981). As can be seen in Table 3, violation of expectations occurred for all groups in the areas of employment and education. For the non-GED group, expectations were not severely violated in relation to personal benefits, while for the two GED groups the results in the personal area were consistent with the results in the other two areas. Table 3 suggests that participants were disappointed in the benefits of the program, expecting much more than they actually received. The trends of disappointment conveyed in the table, even though based on "prior expectations" which were assessed after the fact, seem undeniable.

To check out these trends, respondents were asked if they had participated either full-time or part-time in an educational program since their AHSEP experience. Twenty-nine of the non-GED group, 20% of the GED-pass group and 14% of the GED-fail group had received or were receiving further education. Fewer than 35% in each group had even applied for admission to some kind of training program (GED-pass: 32%; GED-fail: 29%; non-GED: 34%).

Employment patterns for the three groups were also determined through the questionnaires. All groups experienced slight increases in paid employment since participating in the program. Unlike the Behal (1983) study in which high levels of unemployment were reported for all GED candidates, the present study found only the GED-fail group to have an unusually high level of unemployment. The responses on employment are shown in Table 4.

If only the unemployed who sought work are considered, the GED-fail group has an unemployment rate that is over twice as great as the other two groups. Moreover, when prior and post comparisons of employment status are made, only the GED-fail group showed an increase in the percentage of those unemployed but

TABLE 3

Comparison of Prior Expectations and Post Program Benefits
(Figures in Table are Percentages of Respondents)

Area of Expectation	GED-pass		GED-fail		non-GED	
	Prior	Post	Prior	Post	Prior	Post
	n=113		n=77		n=133	
Help in getting a job	43	11*	53	4*	33	8*
Help in a personal way	40	24*	33	7*	29	23
Help in being accepted into an educational institution	38	12*	40	12*	38	16*
Make no differences to job situation	20	46*	8	34*	22	35*
Help in getting a promotion	15	4*	17	1*	21	1*
Help in being accepted into an apprenticeship program	13	2*	29	1*	18	4*
Help in keeping present job	5	4	1	3	8	3
Help in getting raise	4	4	0	1	3	3
Make no difference to educational opportunities	4	11	5	16*	4	6
Not help in any way	2	21*	0	39*	5	20*

* change of more than 10%

TABLE 4

Employment Status
(Figures in Table are Percentages)

	GED-pass		GED-fail		non-GED	
	n=133		n=77		n=133	
	Prior	Post	Prior	Post	Prior	Post
Employed						
Full-time for pay	53	58	36	43	55	58
Part-time for pay	13	12	12	8	11	13
Unemployed						
Not looking for work	8	3	16	7	9	7
Looking for work	16	15	23	33	18	14

looking for work. Such a prior-post finding for the GED-fail group is not surprising for at least two reasons. First, the GED-fail group tended to be younger than either of the other groups and unemployment rates have traditionally been highest, according to Statistics Canada, among young people. This youngest group also has less job market experience, an important employment factor, as Dhanidina and Griffith (1975) have pointed out. Second, according to Statistics Canada, general levels of unemployment in Alberta gradually rose from 4.5% in December 1981 to 11.1% in December 1983, the time period covered by the study.

Levels of unemployment for the GED-pass and non-GED groups declined slightly. Thus, at a time when Alberta unemployment levels were increasing, the unemployment rates of some equivalency program participants actually dropped. It may be that acquiring an equivalency diploma helped some individuals to find new employment. It is also possible that the shifts in employment data for these groups may reflect a more general motivational factor rather than a causal factor.

Exploratory analyses of the employment data, however, suggested no statistically significant differences between the prior and post employment status of any of the three groups [GED-pass (X^2[25, n= 113]=167.2, p<.01); GED-fail (X^2[16, n= 77]=54.8, p <.01); non-GED (X^2[20, n= 133]=208.9, p<.01)]. These findings suggest that there was no change in the employment status of the GED-fail group, in spite of the observed percentage increase of those in this group who were unemployed but looking for work.

Personal benefits outweighed educational and job-related benefits according to the AHSEP participants. Self-satisfaction was the most frequently reported personal benefit. Positive comments concerning the equivalency program by even those who had failed the GED tests focused on personal benefits rather than educational or job-related benefits.

Relationships Between GED Test Performance and Selected Demographic Variables

As reported earlier, preliminary analyses of the survey data revealed that the GED-pass and GED-fail groups could be distinguished on the basis of age but not on the basis of previous highest level of education. To gain a more complete picture of the relationship of these variables to GED test performance, multivariate analyses of respondents' test scores were conducted for both the age and highest-grade-completed variables.

Table 5 provides a breakdown of GED sub-test means for four age groups (19-25, 26-30, 31-40, 41 and older) of GED candidates. The ordering of sub-test means is consistent with the finding that most GED failures in this study were from the youngest age category.

A one-way multivariate analysis of variance (MANOVA) of the vectors of the five GED sub-test means for the four age groups also supports this finding. The one-way MANOVA resulted in an overall significant difference among the four groups (Wilk's Lambda = 0.839, Rao's approximate F (15, 502.8)=2.20, p<.01). Contrasts of the vectors of sub-test means revealed that the scores of the 19 to 25 year-old group were significantly different (p<.05) from those of the other age groups. No other age group differences were found.

A similar analysis using four grade-level groups (grades 9, 10, 11, 12) resulted in no significant differences (Wilk's Lambda = 0.874, Rao's approximate F (15, 502.8)=1.68, p<.05). Inspection of the sub-test means for males and females revealed very small differences (usually less than 0.5 standard score points) and suggested that further multivariate analysis was not appropriate.

TABLE 5

Age Group	GED Sub-test Means[1]				
	Writing	Social Studies	Science	Reading	Math
25 or less (n=69)	48.4	48.6	50.6	51.2	48.0
26 to 31 (n=39)	54.6	57.5	56.8	57.6	53.0
31 to 40 (n=51)	52.7	55.9	55.8	57.5	53.1
Over 40 (n=31)	51.5	54.9	53.7	54.8	52.4

[1]Scores reported in the table are standard T- scores ($u=50$, $o=10$).

Discussion

Given that the goals and objectives for the Alberta High School Equivalency Program are not very different from those of other equivalency programs, it was not surprising that the characteristics of the Alberta equivalency population and other equivalency populations described in the literature were found to be similar. The reasons given for applying to the equivalency progam, prior expectations, and benefits reported by the sample of AHSEP participants were not unlike those of other equivalency programs.

It is evident that individuals who have the most difficulty in passing the GED tests are young, have low levels of formal education, and are often unemployed. For the Alberta equivalency population, age seems to be related to GED test performance, a finding not supported in the literature (Modjeski & Michael, 1978). Also, GED failures are most often motivated by the belief that a credential will help them to get a job or to go on to some form of post-secondary education in order to improve their job situation.

The differential relationship between age and reasons for participation in the equivalency program is an interesting one. Older individuals in this study were motivated to participate in the equivalency program because, generally, they saw it as being personally rewarding. Younger participants, on the other hand, were more pragmatic; upon completion of the program, they hoped to get a better job. This finding reinforces Cross's (1981) contention that the reasons adults participate in education programs "correspond consistently and logically to the life situations" (p. 91) of those individuals. Based on the findings of the present study, more extensive investigations of the relationship of age and motivational factors to GED test performance are in order.

As in similar programs, there was a tendency to inflate the value of the equivalency diploma in terms of its utility in the educational and job marketplaces. Many participants saw the equivalency credential as a way of overcoming all barriers in their attempts to find employment or acceptance to post-secondary education. This may be one reason why discrepancies between prior expectations and post-program benefits, similar to those reported by Cervero and Peterson (1982), were found. The fact that the employment situations of both the GED-pass and non-GED groups remained stable even at a time when unemployment was on the rise, has not tempered the disappointment in the benefits of the program for individuals in these groups.

It appears reasonable to suggest that retrospective judgments of personal benefits outweighed educational or job-related benefits. Self-satisfaction was the most frequently reported personal benefit. This finding is related to but somewhat different from findings reported in the literature. Enhanced self-image or self-confidence is the most frequently mentioned personal benefit of equivalency program participation mentioned in the literature (Cervero & Peterson, 1982; Fischer & Evanson, 1979; Harris et al., 1983; Waller, 1976). The discrepancy between the present study and those just cited may have arisen as a result of different data collection procedures. GED surveys most often directly ask candidates if they feel that their self-confidence or self-image has improved as a result of the GED experience thereby increasing the likelihood of identifying this benefit. In the present study, participants were free to volunteer how they believed the equivalency program had helped them. The importance of self-concept was evident for the Alberta equivalency program in the same way as it has been demonstrated to be an important component of adult education programs in general by Wilson (1979; 1980) and Rolfe and Wilson (1979).

While the extent to which the Alberta High School Equivalency Program assists individuals in meeting their vocational and personal educational goals must be questioned, the value of personal benefits received by participants should not be underestimated. Psychological benefits may have assisted some of the participants in finding a new job or may have provided the impetus for others to apply to a post-secondary institution. It may be that the credential itself is less important than the positive outlook created by the success in acquiring it.

Note

1. This research was supported in part by a grant from the Alberta Advisory Committee for Educational Studies.

References

American Council of Education (1982). *College admissions policies toward GED applicants.* Washington, DC: Author.

Ayers, C.V. (1980). *An educational and employment growth study of General Educational Development (GED) graduates of Surrey Community College.* (ERIC Document Reproduction Service No. ED 196 472).

Behal, E.L. (1983). *After the GED tests: Post-secondary education enrollment patterns and perceptions of GED examinees* (GED Testing Service Research Studies, No. 5, September 1983). Washington, DC: American Council on Education.

Beusse, W.E. (1974). In-service and post-service benefits of General Educational Development (GED) program participation. *Catalog of Selected Documents in Psychology, 4* (101), MS 722.

Cervero, R.M., & Peterson, N.L. (1982). *After the GED tests: Employment and training activities of GED examinees* (Ged Testing Service Research Studies, Number 4, November, 1982). Washington, DC: American Council on Education.

Cross, K.P. (1981). *Adults as learners.* San Francisco: Jossey-Bass Publishers.

Dhanidina, L., & Griffith, W.S. (1975). Costs and benefits of delayed high school completion. *Adult Education, 25*(4), 217-230.

Fischer, J.K., & Evanson, J.I. (1979). Test scores don't tell the whole story. *New Directions for Continuing Education, 3,* 29-36.

Harris, I., Hedman, C., & Horning, M. (1983). Success with high school dropouts. *Educational Leadership, 40*(3), 35-36.

Harvey, J.R. (1978). Personal and motivational characteristics of adult learners as related to Houle's typology. *Dissertation Abstracts International, 39*, 1992A-1993A.

Laurence, J.H. (1983, April). *Educational credentials and military enlistment*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec.

Malizio, A.G., & Whitney, D.R. (1981). *Who takes the GED tests? A national survey of spring 1980 examinees* (GED Testing Service Research Studies, No. 1, March, 1981). Washington, DC: American Council on Education.

Malizio, A.G., & Whitney, D.R. (1985). *Educational credentials in employment: A nationwide survey* (Office on Educational Credit and Credentials, May, 1985). Washington, DC: American Council on Education.

Manitoba Department of Education (1981). *GED acceptance with employers—1980-81*. Winnipeg, MB: Author.

Manitoba Department of Education (1982). *High school equivalency follow-up survey report*. Winnipeg, MB: Author.

Modjeski, R.B., & Michael, W.B. (1978). The relationship of the General Educational Performance Index measure to other indicators of educational development in each of three samples from an United States Army population. *Educational and Psychological Measurement, 38* (2), 377-391.

Moore, L. (1982). *A follow-up study of adults who were administered the General Educational Development test in Kentucky during 1976*. Frankfort, KY: Department of Education. (ERIC Document Reproduction Service No. ED 222 583.)

Province of Saskatchewan (1977). *GED acceptance survey*. Regina, SK: Author.

Province of Saskatchewan (1981). *Acceptance of General Educational Development (GED) certification by Saskatchewan employers and post-secondary educational institutes*. Regina, SK: Author.

Province of Saskatchewan (1982). *Follow-up survey of 1980 GED graduates*. Regina, SK: Author.

Rolfe, P.G., & Wilson, R.C. (1979). *A comparison of the self-descriptions of high school completion (GED) dropouts and persistors*. Auburn, AL: Auburn University, Department of Vocational and Adult Education. (ERIC Document Reproduction Service No. Ed. 186 776.)

Russ-Eft, D.R., & Steel, L.M. (1980). Contribution of education to adults' quality of life. *Educational Gerontology, 5*, 189-209.

Statistics Canada (1980). *Perspectives Canada III*. Ottawa, ON: Supply and Services Canada.

Swarm, C. (1981). *Three studies of General Educational Development students—1971-1981*. (ERIC Document Reproduction Service No. ED 211 696.)

Waller, E.A. (1976). *Summary of research findings on the Military General Educational Development Program* (Report No. AFHRL TR 76-55). Brooks Air Force Base, TX: Air Force Human Resources Laboratory (ERIC Document Reproduction Service No. ED 139 986).

Wedel, G.J. (1974). The General Educational Development high school equivalency diploma in the Province of Saskatchewan (Doctoral dissertation, University of Northern Colorado, 1974). *Dissertation Abstracts International, 35* 6953A-6954A.

Wilson, R.C. (1979). Personological profile of community college high school equivalency students. *Community College Review, 6*, 52-59.

Wilson, R.C. (1980). Personological variables related to GED retention and withdrawal. *Adult Education, 30*, 173-185.

Witter, E. (1978). GED teachers are a new breed. *Phi Delta Kappan, 60* (2), 110-111.

The Alberta Journal of Educational Research *Vol. XXXII, No. 1, March, 1986, 77-78*

BOOK REVIEWS

CURRICULUM CANADA V. *Edited by Douglas A. Roberts and John O. Fritz. Vancouver, B.C: Centre for the Study of Curriculum and Instruction, University of British Columbia, 1984, 215 pp.*

Every year, the Canadian Association for Curriculum Studies sponsors an invitational conference dealing with curriculum concerns in the Canadian context. The papers presented at these conferences are published by the Centre for the Study of Curriculum and Instruction at the University of British Columbia. *Curriculum Canada V* is the fifth in the series.

It is worth noting at the outset that these volumes, taken together, provide a valuable resource for scholars in the field of Curriculum Studies. The papers not only address current issues in the field but do so from a Canadian perspective. This perspective had been noticeably missing from the literature in previous years.

When a volume to be reviewed consists of conference papers, one must ask: How well did the papers achieve the purpose of the conference? That is the question which will be addressed in this review.

The conference resulting in *Curriculum Canada V* recognized that there are two groups of stakeholders in the field of Curriculum Studies: subject area specialists (e.g., reading, mathematics) and generalists whose concerns cross subject matter boundaries. As the editors of this volume write: "Clearly the relationship between the two needs exploration." Therefore, they explain, "The major objective of the conference was to clarify issues involved in relating the school subject area research to more general curriculum/instruction issues."

The conference was organized around presenters representing six major school subjects: science, reading, social studies, language arts, mathematics, and second languages. The presenters approached the conference theme through a "state of the art" appraisal in each of the six areas. Each presenter prepared a paper in advance of the conference, and a response was prepared by another scholar in the same subject area.

The conference began with a keynote address by Dr. George Tomkins on the history of curriculum in Canada. This paper provides an excellent overview of the trends which have brought curriculum in Canada to its present state. Tomkins' paper says basically "Here's where we've been." The other papers go on to say "Here's where we are now."

The papers devoted to specific subject areas describe current research in those subject areas as well as the issues involved. This is accomplished in an interesting and informative manner. The respondents' papers broaden the reader's thinking either by agreeing with the first author and providing additional support or by disagreeing and presenting the reader with another point of view. Both approaches prove to be valuable.

However, the presenters were also expected to "address themselves to issues of general concern for curriculum/instruction thought." It is this link between subject area concerns and broader curriculum concerns which is less evident in the papers. Although some of the authors touched on general curriculum concerns, they seldom identified them as such. A notable exception is the paper by Ken Osborne which concludes with a section on "Social Studies and General Curriculum Theory."

One wonders if the presenters were in a position to relate the issues in their subject areas to "general" curriculum theory. They are undoubtedly experts in their subject areas, but they are not necessarily experts in "general" curriculum theory as well. Perhaps they were being asked to do something which they were not equipped to do.

The problem may lie, not with the presenters, but with the organization of the conference. In order to bridge the gap between subject area specialists and general curriculum specialists, representatives of both groups would need to talk together. It is interesting, then, that the presenters and respondents were all subject area specialists. Members of the generalist group were apparently not asked to participate.

Their participation, however, would have been useful. For example, if the respondents to the subject-oriented papers had been generalists, the focus of their remarks could have been "Here's how general curriculum research can contribute to your field" or "Here's how the research which you have described in your field can contribute to general curriculum theory." Then a meeting of the minds could have been broached between the two groups.

As it is, *Curriculum Canada V* is a useful volume from the standpoint of current research in the various subject areas. The state of each field is presented well. However, a discussion between subject area specialists and general curriculum specialists is not to be found in this particular collection of papers. It is hoped that a future conference will initiate such a dialogue.

Jean H. Young
Department of Elementary Education
University of Alberta

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in triplicate and are not to exceed 25 pages in length.
2. An abstract of approximately 100 words in length, typed on a separate page, is to be included.
3. The author's name and affiliation should appear only on a separate title page to ensure anonymity in the reviewing process; the title of the paper should appear on the first page of the manuscript as a means of identification.
4. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publication Manual.
5. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
6. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
7. Sources should appear in parentheses following the reference, citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
8. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 7 above.
9. Spelling shall conform to *Webster's New Collegiate Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
10. In matters of style, the third edition of the APA Publication Manual is considered definitive.

The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*